L. Pandolfi

Corso di Analisi Funzionale (versione 2022)

Questa versione è scritta in carattere romano e giustificata a destra. Una versione più leggibile da chi ha problemi di dislessia (scritta in carattere bastone e non giustificata) ha per titolo Corso di Analisi Funzionale (versione 2022) Per Sudenti Dislessici.

L. Pandolfi, in pensione dal Politecnico di Torino, Dipartimento di Scienze Matematiche "G.L. Lagrange"

Le abbreviazioni usate in questo testo sono di immediata comprensione:

- “s.successione” per “sottosuccessione”
- “s.insiemi” per “sottoinsiemi”
- “s.insieme” oer “sottinsieme”
- “s.spazio” per “sottospazio”
- “s.spazi” per “sottospazi”
- “s.l.n.” per “spazio lineare normato”
- “s.l.n-ti” per “spazi lineari normati”.
- “spazio lineare” spazio lineare ”spazi lineari” spazi lineari

Indice

Capitolo 1

Introduzione all'analisi funzionale

1.1 Equazioni integrali di Fredholm ed analisi funzionale

L'*Analisi funzionale* è nata quando, tra la fine del XIX e il primo trentennio del XX secolo, sono stati raccolti in un'unica teoria molti risultati provenienti da varie parti dell'analisi matematica; in particolare, dalla teoria delle EQUAZIONI INTEGRALI della forma

$$x(s) = \mu \int_a^b K(s,\xi)x(\xi)\,\mathrm{d}\xi + y(s)\,, \qquad s \in [a,b] \tag{1.1}$$

ove μ è un parametro e $K(s,\xi)$, $y(s)$ sono funzioni note mentre la funzione $x(s)$ è incognita. L'intervallo $[a,b]$ è *limitato.*

Equazioni di questo tipo si chiamano EQUAZIONI DI FREDHOLM DI SECONDA SPECIE e si incontrano per esempio nella soluzione di problemi di elasticità (problemi che hanno stimolato gli studi su tali equazioni).

La funzione $K(s,\xi)$ si chiama il NUCLEO dell'equazione di Fredholm.

Si chiama EQUAZIONE DI FREDHOLM DI PRIMA SPECIE l'equazione

$$\int_a^b K(s,\xi)x(\xi)\,\mathrm{d}\xi = y(s)\,, \qquad s \in [a,b] \tag{1.2}$$

che pure si incontra nelle applicazioni, ma che ha una teoria sostanzialmente più delicata, per una ragione che vedremo.

Come introduzione all'analisi funzionale, vediamo come sia possibile trattare l'equazione di Fredholm (di prima o di seconda specie) quando il nucleo

è DEGENERE,indexnucleo!dell'equazione di Fredholm!degenere ossia quando

$$K(s,\xi) = \sum_{i=1}^{m} a_i(s)b_i(\xi)\,, \qquad s\,,\ \xi \in [a,b]\,. \tag{1.3}$$

Per semplicità supporremo che *le funzioni* $a(s)$ *e* $b(\xi)$ *siano continue su* $[a,b]$.

Nel caso del nucleo degenere la soluzione dell'equazione di Fredholm si riduce ad un problema di algebra lineare e quindi richiamiamo alcune nozioni che dovremo usare.

1.1.1 L'equazione $Ax = \phi$

Ricordiamo che $\mathbb{C}$ indica l'insieme dei numeri complessi e che $\mathbb{C}^n$ indica il prodotto cartesiano di n copie di $\mathbb{C}$. In $\mathbb{C}^n$ esistono infinite basi, ciascuna di n elementi. Si chiama BASE CANONICA la base $\{e_1\,,\ldots\,,e_n\}$, avendo indicato con e_i il vettore le cui componenti sono tutte nulle, salvo la i–ma, che vale 1.

Se x è un vettore di $\mathbb{C}^n$, scriveremo

$$x = \mathrm{col}\,\left[\begin{array}{ccc} x_1 & \ldots & x_n \end{array}\right] = \sum_{i=1}^{n} x_i e_i\,.$$

Affermazioni del tutto analoghe valgono per $\mathbb{C}^m$. Per evitare ambiguità, indicheremo con $\{\epsilon_1\,,\ldots\,,\epsilon_m\}$ la sua base canonica. Anzi, per chiarezza indicheremo con lettere romane i vettori di $\mathbb{C}^n$ e con lettere greche quelli di $\mathbb{C}^m$.

Sia A una matrice $m \times n$,

$$A = \left[\begin{array}{cccc} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{array}\right]\,.$$

L'espressione $Ax = \phi$ si usa per rappresentare in modo compatto il sistema di m equazioni in n incognite

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & \phi_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & \phi_2 \\ \vdots & & & & & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & \phi_m\,. \end{array} \tag{1.4}$$

Vogliamo ricordare le condizioni, note dai corsi precedenti, perché:

1. L'equazione $Ax = \phi$ ammetta *al più una soluzione;*
2. L'equazione $Ax = \phi$ ammetta *almeno una soluzione;*
3. L'equazione $Ax = \phi$ ammetta *esattamente una soluzione per ogni* $\phi \in \mathbb{C}^m$*.*

Consideriamo prima di tutto il problema 1), *unicità di soluzione.* Si noti la formulazione del problema di unicità: *non si richiede che una soluzione debba esistere* per ogni ϕ; *Si richiede che* se una soluzione esiste *allora questa sia* unica.

Se per la medesima ϕ si hanno due soluzioni x' ed x'' tra loro diverse, allora $y = x' - x''$ verifica

$$Ay = A(x' - x'') = Ax' - Ax'' = \phi - \phi = 0\,.$$

Dunque, se *per una* ϕ la soluzione non è unica, allora esiste $y \neq 0$ e tale che $Ay = 0$. E viceversa:

Teorema 1 *Se per* una ϕ *il problema (1.4) ammette soluzione e questa è unica, allora*

$$Ax = 0 \text{ solo se } x = 0.$$

Viceversa, se Ax si annulla solo per $x = 0$ allora il problema (1.4) ammette soluzione unica per ogni ϕ*.*

Osservazione 2 Si noti che il problema dell'unicità, come è stato enunciato, sembra dipendere dalla scelta di ϕ. Ossia, si potrebbe pensare che la soluzione di (1.4) sia unica per alcune ϕ ma non per altre. Invece, il teorema precedente mostra che si ha unicità di soluzione *per ogni* ϕ se e solo se si ha unicità di soluzione *per una* ϕ*.* ∎

L'insieme

$$\ker A = \{x \mid Ax = 0\}$$

si chiama il NUCLEO di A.

Passiamo ora a studiare il *problema dell'esistenza di soluzioni.* Chiamiamo IMMAGINE di A l'insieme

$$\operatorname{im} A = \{Ax \mid x \in \mathbb{C}^n\}\,.$$

Dunque, $\operatorname{im} A$ è esattamente l'insieme dei vettori ϕ per i quali l'equazione (1.4) è risolubile.

Si ricordi che $\operatorname{im} A$ è un s.spazio di $\mathbb{C}^m$. Ora un generico s.spazio V di $\mathbb{C}^m$ può caratterizzarsi descrivendone gli elementi; e ciò abbiamo fatto per definire $\operatorname{im} A$; ma può anche identificarsi per mezzo del suo s.spazio ORTOGONALE $V^\perp$.

Definizione 3 Il prodotto interno dei due vettori $\psi = \sum_{i=1}^m \psi_1\epsilon_i$ e $\phi = \sum_{i=1}^m \phi_1\epsilon_i$ di $\mathbb{C}^m$ è

$$\langle \phi, \psi \rangle = \sum_{i=1}^m \bar{\psi}_n \phi_n \,.$$

I due vettori si dicono ortogonali quando il loro prodotto interno è nullo:

$$\langle \phi, \psi \rangle = \sum_{i=1}^m \bar{\psi}_n \phi_n = 0\,.$$

Un vettore $\psi \in \mathbb{C}^m$ si dice ortogonale a $V \subseteq \mathbb{C}^m$ quando è ortogonale a tutti gli elementi di V; e $V^\perp$ è l'insieme di tutti i vettori ψ ortogonali a V.

Definizioni analoghe valgono anche in $\mathbb{C}^n$. ∎

Dunque,

$$V^\perp = \{\psi \mid \langle \phi, \psi \rangle = 0 \; \forall \phi \in V\}\,.$$

Si vede facilmente che:

- *$V^\perp$ è un s.spazio anche se V non è un s.spazio.*
- *vale $V^\perp = 0$ se e solo se V è uguale a $\mathbb{C}^m$.*

Quando $V = \operatorname{im} A$, è facile identificare $V^\perp$. Introduciamo per questo la matrice A^*, aggiunta di A. Si sa che A^* verifica

$$\langle Ax, \phi \rangle = \langle x, A^*\phi \rangle\,. \tag{1.5}$$

Teorema 4 *si ha*

$$(\operatorname{im} A)^\perp = \ker A^*\,, \qquad \operatorname{im} A = (\ker A^*)^\perp\,.$$

<u>Dim.</u> Se $\phi \perp \operatorname{im} A$ allora *per ogni x* vale

$$\langle Ax, \phi \rangle = 0 \qquad \text{ossia} \qquad \langle x, A^*\phi \rangle = 0 \quad \forall x \in \mathbb{C}^n\,;$$

e quindi $A^*\phi = 0$. Viceversa, se $A^*\phi = 0$ allora *per ogni $x \in \mathbb{C}^n$* vale

$$0 = \langle x, A^*\phi \rangle = \langle Ax, \phi \rangle \qquad \forall x \in \mathbb{C}^n$$

e quindi $\phi \perp \operatorname{im} A$. Ciò prova la prima uguaglianza. La seconda si può provare in modo analogo, ma può anche dedursi dalla prima, ricordando che $(A^*)^* = A$ e che $\left(V^\perp\right)^\perp = V$ per ogni s.spazio di $\mathbb{C}^m$. ∎

Usando ancora le uguaglianze $(A^*)^* = A$ e $(V^\perp)^\perp = V$ si trova:

Corollario 5 *Vale:*

$$\ker A = (\operatorname{im} A^*)^\perp\,, \qquad \operatorname{im} A^* = (\ker A)^\perp\,.$$

Dunque, $\ker A = 0$ *se e solo se* $(\operatorname{im} A^*)^\perp = 0$*, ossia se* A^* *è suriettiva.*

Torniamo ora all'equazione (1.4). I risultati precedenti mostrano che:

Teorema 6 *L'equazione (1.4) è risolubile se e solo se* $\phi \perp (\ker A^*)\,.$

Questo teorema assume un aspetto particolarmente semplice se vogliamo che l'equazione (1.4) sia risolubile *per ogni* ϕ*:*

Teorema 7 *L'equazione (1.4) è risolubile* per ogni ϕ *se e solo se* $\ker A^* = \{0\}$.

Si può dunque enunciare:

Teorema 8 (ALTERNATIVA DI FREDHOLM) *La matrice A identifica una trasformazione* iniettiva *se e solo se la matrice* A^* *identifica una trasformazione* suriettiva; *La matrice A identifica una trasformazione* suriettiva *se e solo se la matrice* A^* *identifica una trasformazione* iniettiva.

Osservazione 9 L'alternativa di Fredholm per le matrici è utile perché in pratica è assai più semplice verificare l'*unicità* di soluzione piuttosto che l'*esistenza* di soluzione. ∎

L'alternativa di Fredholm si incontra anche in situazioni più generali dello studio di sistemi di n equazioni in m incognite. Però, nel caso particolare del problema (1.4) si può essere anche più precisi e notare che se $\ker A = \{0\}$ e inoltre A è suriettiva, allora deve essere $n = m$.

Si ricordi che una trasformazione A da $\mathbb{R}^n$ in sé è suriettiva ed iniettiva quando, fissata una base, è diverso da zero il determinante della matrice ad essa corrispondente.

1.1.2 L'equazione $\lambda x - Ax = y$

Sia ora $n = m$ e quindi A sia una matrice quadrata, e studiamo l'equazione

$$\lambda x - Ax = y\,, \qquad \text{ossia} \quad (\lambda I - A)x = y \tag{1.6}$$

ove I indica la matrice identità. Naturalmente quest'equazione è un caso particolare di (1.4) e quindi si tratta di adattare a questo caso particolare i risultati già trovati. In particolare:

Teorema 10 *L'equazione (1.6) ammette* una e una sola *una soluzione* per ogni y *se e solo se l'equazione*

$$(\bar{\lambda}I - A^*)\psi = \phi$$

ammette esattamente una soluzione. Ciò avviene se e solo se

$$\det(\lambda I - A) \neq 0\,. \tag{1.7}$$

L'insieme dei numeri λ per cui vale l'unicità di soluzione si chiama L'INSIEME RISOLVENTE della matrice A e si indica col simbolo $\rho(A)$; il suo complementare si chiama lo SPETTRO della matrice e si indica col simbolo $\sigma(A)$. Gli elementi di $\sigma(A)$ si chiamano[1] gli AUTOVALORI di A.

Si noti che

$$\det(\lambda I - A)$$

è un polinomio di grado $n \geq 1$; e quindi *lo spettro di una matrice non è vuoto.*

Gli autovalori sono in generale numeri complessi, anche se la matrice A è reale. Si ricordi che se la matrice A è reale allora $\lambda \in \sigma(A)$ se e solo se $\bar{\lambda} \in \sigma(A)$.

In generale si ha che se $\lambda \in \sigma(A)$ allora $\bar{\lambda} \in \sigma(A^*)$.

Si lascia per esercizio di enunciare risultati analoghi per l'equazione

$$x = \mu Ax + y\,. \tag{1.8}$$

Notiamo però che quest'equazione è risolubile in modo banale quando $\mu = 0$ perché in tal caso essa si riduce a $x = y$.

1.1.3 L'equazione di Fredholm di seconda specie con nucleo degenere

Studiamo ora l'equazione (1.1) nel caso particolare in cui il nucleo ha forma (1.3), ossia nel caso del NUCLEO DEGENERE.

Il parametro μ moltiplica l'integrale, così che l'equazione di Fredholm così scritta viene ad essere di seconda specie per ogni valore di μ. Avessimo invece scritto

$$\lambda x(s) = \int_a^b K(s,\xi)x(\xi)\,\mathrm{d}\xi + y(s)\,, \qquad s \in [a,b]$$

per $\lambda = 0$ avremmo trovato un'equazione di prima specie.

[1] vedremo che nel caso più generale di trasformazioni tra spazi di funzioni gli autovalori sono solo una parte dello spettro.

L'equazione (1.1) si scrive

$$x(s) = \mu \sum_{r=1}^{m} a_r(s) \int_a^b b_r(\xi) x(\xi)\, \mathrm{d}\xi + y(s) \tag{1.9}$$

ossia

$$x(s) = \mu \sum_{r=1}^{m} a_r(s) x_r + y(s)\,, \qquad x_r = \int_a^b b_r(\xi) x(\xi)\, \mathrm{d}\xi\,.$$

I numeri x_r dipendono dalla funzione incognita $x(s)$ e quindi non sono noti; però, moltiplicando i due membri dell'uguaglianza per $b_i(s)$ e integrando da a a b, si trova che essi risolvono

$$x_i = \mu \sum_{r=1}^{m} k_{ir} x_r + y_i\,, \qquad \begin{cases} k_{ir} &= \int_a^b b_i(s) a_r(s)\, \mathrm{d}s \\ y_i &= \int_a^b b_i(s) y(s)\, \mathrm{d}s\,. \end{cases} \tag{1.10}$$

Dunque, i numeri x_i risolvono un sistema della forma (1.8). Introduciamo allora la matrice K i cui elementi sono k_{ir} e i vettori x e y, i cui elementi sono rispettivamente i numeri x_i e y_i e scriviamo il sistema (1.10) come

$$(I - \mu K)x = y\,. \tag{1.11}$$

Dunque ogni soluzione dell'equazione di Fredholm (1.9) identifica una soluzione x di (1.11) e si vede anche che vale il viceversa: se x risolve (1.11) allora

$$x(s) = \mu \sum_{r=1}^{m} a_r(s) x_r + y(s)$$

risolve l'equazione (1.9). Mostriamo infatti che questa funzione, sostituita a destra ed a sinistra di (1.9), verifica l'uguaglianza. A sinistra si trova

$$\mu \sum_{r=1}^{m} a_r(s) x_r + y(s)$$

mentre sostituendo a destra e tenendo conto di (1.10) (e scambiando il nome degli indici) si trova

$$\begin{aligned} &\mu \sum_{r=1}^{m} a_r(s) \int_a^b b_r(\xi) \left\{ \mu \sum_{i=1}^{m} a_i(\xi) x_i + y(\xi) \right\} \mathrm{d}\xi + y(s) \\ &= \mu \sum_{r=1}^{m} a_r(s) \left\{ \mu \sum_{i=1}^{m} k_{ri} x_i + y_r \right\} + y(s) = \mu \sum_{r=1}^{m} a_r(s) x_r + y(s)\,, \end{aligned}$$

come si voleva.

Possiamo quindi trasferire all'equazione integrale di Fredholm i risultati che abbiamo enunciato per i sistemi di equazioni lineari:

Teorema 11 *L'equazione di Fredholm (1.1) con nucleo degenere ammette al più una soluzione se e solo se l'equazione*

$$\psi = \bar{\mu} K^* \psi + \phi \tag{1.12}$$

è risolubile per ogni ϕ; l'equazione di Fredholm con nucleo degenere è risolubile per ogni funzione $y(x)$ se e solo se l'equazione (1.12) ammette unicità di soluzione.

Il teorema precedente dà una specie di "alternativa di Fredholm" per l'equazione integrale, ma trascritta mediante le matrici. E' però facile vedere che l'equazione (1.12) corrisponde all'equazione integrale

$$\psi(\xi) = \bar{\mu} \int_a^b \overline{K(\xi, s)} \psi(\xi) \, \mathrm{d}\xi + \phi(\xi) \,, \tag{1.13}$$

equazione che si chiama AGGIUNTA della (1.1). Possiamo quindi enunciare l'ALTERNATIVA DI FREDHOLM per le equazioni integrali di Fredholm di seconda specie a nucleo degenere:

Teorema 12 *Quando il nucleo è degenere, la (1.1) ammette unicità di soluzione se la sua aggiunta (1.13) è risolubile per ogni $\phi(s)$; la (1.1) è risolubile per ogni $y(s)$ se la sua aggiunta (1.13) ammette unicità di soluzione.*

Ancora, è assai più facile verificare l'unicità piuttosto che l'esistenza di soluzioni, e ciò spiega l'interesse di questo teorema.

1.1.4 L'equazione di Fredholm di prima specie

Le considerazioni precedenti possono tutte ripetersi nel caso dell'equazione di prima specie (1.2). In tal caso, invece della (1.10) si trova l'equazione

$$\sum_{r=1}^{m} k_{ir} x_r = y_i$$

ossia

$$Kx = y \,. \tag{1.14}$$

Per vedere la differenza tra questa e l'equazione (1.11), esaminiamo il caso particolare in cui

$$k_{ir} = \int_a^b b_i(\xi) a_r(\xi)\, \mathrm{d}\xi = \begin{cases} 0 & \text{se} \quad i \neq r \\ k_{ii} \neq 0 & \text{se} \quad i = r\,. \end{cases}$$

In tal caso la (1.11) e la (1.14) divengono rispettivamente

$$\begin{bmatrix} (1-\mu k_{11}) & & & \\ & (1-\mu k_{22}) & & \\ & & \ddots & \\ & & & (1-\mu k_{mm}) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix},$$

$$\begin{bmatrix} k_{11} & & & \\ & k_{22} & & \\ & & \ddots & \\ & & & k_{mm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

Ambedue queste equazioni sono, *in principio,* facilmente risolubili e le soluzioni sono, rispettivamente,

$$\begin{bmatrix} y_1/[1-\mu k_{11}] \\ y_2/[1-\mu k_{22}] \\ \vdots \\ y_m/[1-\mu k_{mm}] \end{bmatrix}, \qquad \begin{bmatrix} y_1/k_{11} \\ y_2/k_{22} \\ \vdots \\ y_m/k_{mm} \end{bmatrix}.$$

In generale, al crescere dell'indice i, i numeri k_{ii} divengono "velocemente" piccoli e quindi nel caso dell'equazione di prima specie, gli errori commessi nella misura di $y(s)$, e quindi nel calcolo degli y_i, si amplificano velocemente. Ciò non accade per il caso dell'equazione di seconda specie, salvo nel caso in cui $1/\mu$ è circa uguale ad uno dei numeri k_{ii}; e anche in questo caso la sola componente i–ma può provocare dei problemi numerici.

Queste considerazioni mostrano che, almeno nel caso del nucleo degenere, la soluzione delle equazioni integrali di prima specie è assai più delicata della soluzione di quelle di seconda specie.

1.1.5 Ricapitolazione

Ora poniamoci alcuni problemi, la cui soluzione guiderà la scelta degli argomenti di analisi funzionale che studieremo. Il primo è questo: tutti i nuclei espressi mediante polinomi sono degeneri. Ma per esempio

$$k(t,s) = e^{s\xi}$$

non è un nucleo degenere. Quindi possiamo chiedere se sia possibili approssimare nuclei abbastanza regolari mediante nuclei degeneri. A questo problema risponderà il *teorema di Weierstrass,* teorema 47.

Per applicare il teorema 12 è necessario ricondursi alla situazione matriciale del teorema 11. In tal caso l'unicità di soluzione si ha quando $1/\mu$ non è un'autovalore della matrice A. Se però il nucleo non è degenere, ciò certamente non può farsi; e allora ci si chiede se l'alternativa di Fredholm, opportunamente reinterpretata, continui a valere e, in caso affermativo, come sia possibile verificare l'unicità di soluzione.

Vorremo infine capire se anche nel caso dei nuclei non degeneri l'equazione di prima specie sia più delicata di quella di seconda, e chiarire la ragione di ciò.

Fatte queste premesse, passiamo ad uno studio sistematico dell'Analisi Funzionale.

Capitolo 2

Gli spazi di Banach

2.1 Spazi lineari normati

Richiamiamo, dai corsi precedenti, le definizioni di spazio lineare e di norma. In questo corso gli spazi lineari avranno sempre, come campo scalare, il campo $\mathbb{R}$ o, più frequentemente, $\mathbb{C}$. Per indicare genericamente uno di questi due campi useremo il simbolo Φ, specificando, se necessario, quando si intende $\Phi = \mathbb{R}$ oppure $\Phi = \mathbb{C}$.

Gli elementi di Φ si chiamano gli SCALARI.

Definizione 13 (di SPAZIO LINEARE) Si dice che X è uno spazio lineare su Φ quando è un gruppo commutativo (rispetto ad un'operazione usualmente indicata col segno $+$) ed inoltre esiste un'operazione (usualmente indicata con notazione moltiplicativa) che ad ogni coppia $(\alpha, x) \in \Phi \times X$ associa un elemento di X e che verifica le seguenti proprietà:

$$\begin{array}{lll} \alpha(x+y) = \alpha x + \alpha y & \forall \alpha \in \Phi\,, & \forall x\,,\ y \in X \\ (\alpha+\beta)x = \alpha x + \beta x & \forall \alpha\,,\ \beta \in \Phi\,, & \forall x \in X \\ 1x = x & & \forall x \in X\,. \ \blacksquare \end{array}$$

Regole di calcolo che discendono dalla precedente definizione sono:

$$\begin{array}{ll} 0x = 0 & \forall x \in X \\ (-\alpha x) = -(\alpha x) & \forall \alpha \in \Phi\,,\ \forall x \in X \\ \alpha 0 = 0 & \forall \alpha \in \Phi\,. \end{array}$$

Ricordiamo alcuni esempi di spazi lineari.

Esempio 14 Gli spazi $\mathbb{R}^n$ e $\mathbb{C}^n$ sono esempi di spazi lineari rispettivamente su $\mathbb{R}$ e su $\mathbb{C}$. (lo spazio $\mathbb{C}^n$ è anche uno spazio lineare su $\mathbb{R}$). Gli spazi dei polinomi, delle funzioni continue, delle funzioni integrabili secondo Riemann; lo spazio delle successioni; o quello delle successioni convergenti; o quello delle successioni limitate; o quello delle successioni convergenti a zero sono sono spazi lineari su $\mathbb{R}$ oppure su $\mathbb{C}$. ∎

Sia X uno spazio lineare su Φ e sia $\mathcal{G}$ un suo s.insieme finito o meno. Si dice che $\mathcal{G}$ è un insieme di GENERATORI se per ogni $x \in X$ esistono un numero naturale n; elementi $g_1, \ldots, g_n$ di $\mathcal{G}$; scalari $\alpha_1, \ldots, \alpha_n$, tali che

$$x = \sum_{i=1}^{n} \alpha_i g_i \,.$$

Se in X esiste un insieme $\mathcal{G}$ *finito* di generatori, si dice che X ha DIMENSIONE FINITA; si dice che ha DIMENSIONE INFINITA altrimenti.

La DIMENSIONE di X è il numero degli elementi di uno dei generatori di X che ha il *minimo* numero di elementi.

Gli spazi $\mathbb{R}^n$ hanno dimensione n; lo spazio $\mathbb{C}^n$ ha dimensione n su $\mathbb{C}$ e dimensione $2n$ su $\mathbb{R}$.

Ha dimensione finita lo spazio dei polinomi di grado al più n. Se come campo scalare si prende il campo cui appartengono i coefficienti, la dimensione è $n+1$. Nonostante questi esempi importanti, in questo corso noi studieremo solamente spazi lineari di dimensione infinita.

Ricordiamo ora la definizione di norma.

Definizione 15 Sia X uno spazio lineare su Φ. Si chiama NORMA su X una funzione, che si indica col simbolo $||x||$, da X in $\mathbb{R}$ (anche quando il campo scalare è $\mathbb{C}$), che soddisfa:

1) $||x|| \geq 0 \qquad \forall x \in X$
2) $||x|| = 0$ se e solo se $x = 0$
3) $||\alpha x|| = |\alpha| \cdot ||x|| \qquad \forall \alpha \in \Phi\,,\ \forall x \in X$
4) $||x + y|| \leq ||x|| + ||y|| \qquad \forall x\,, y \in X\,.$ ∎

L'ultima proprietà della norma si chiama DISUGUAGLIANZA TRIANGOLARE.

Quando c'è possibilità di confusione, si scrive $|| \cdot ||_X$ per la norma dello spazio X.

Uno spazio lineare dotato di norma si chiama uno SPAZIO LINEARE NORMATO e da ora in poi scriveremo semplicemente s.l.n.

Osservazione 16 E' curioso notare che la definizione di norma è ridondante. Ovviamente, la terza proprietà con $\alpha = 0$ implica che

$$||0x|| = |0| \cdot ||x|| = 0$$

e $0x = 0$, l'elemento nullo di X. Dunque una delle due implicazioni in **2)** discende da **3)**.

Anche la proprietà **1)** discende da **2)** e da **3)**. Infatti, essendo $||0|| = 0$ si ha anche, per ogni $x \in X$,

$$0 = ||x - x|| \leq ||x|| + ||-x|| = ||x|| + |-1| \cdot ||x|| = 2||x|| \,. \quad \blacksquare$$

Se X è uno s.l.n., la funzione

$$d(x, y) = ||x - y||$$

è una DISTANZA per la quale

$$d(x + z, y + z) = d(x, y)\,;$$

ossia, è una distanza INVARIANTE PER TRASLAZIONI.

L'introduzione di una norma su X permette di definire:

1. gli INTORNI $B(x_0, \epsilon) = \{x \mid ||x - x_0|| < \epsilon\}$, con $\epsilon > 0$;

2. gli INSIEMI APERTI di X, e quindi una TOPOLOGIA su X: $A \subseteq X$ è aperto se esso contiene un intorno di ciascun suo punto;

3. gli insiemi LIMITATI: un insieme si dice limitato quando è contenuto in un intorno di 0;

4. la CONVERGENZA di successioni: si dice che la successione (x_n) converge ad x_0,
$$\lim x_n = x_0\,,$$
se per ogni $\epsilon > 0$ esiste N_ϵ tale che da $n > N_\epsilon$ segue
$$||x_n - x_0|| < \epsilon \qquad \text{ossia } x_n \in B(x_0, \epsilon).$$

In modo del tutto analogo si definiscono limiti e continuità di funzioni che operano tra s.l.n-ti: per esempio $f(\cdot)$ da X in Y è CONTINUA nel punto $x_0 \in X$ se è ivi definita e se per ogni $\epsilon > 0$ esiste $\delta > 0$ tale che

$$||x - x_0||_X < \delta\,,\ x \in \mathrm{dom} f(\cdot)\,, \quad \text{segue} \quad ||f(x) - f(x_0)||_Y < \epsilon\,.$$

Esattamente come nel caso dei valori assoluti e dei moduli dei numeri complessi, si prova:

Teorema 17 *Vale:*

$$\Big| ||x|| - ||y|| \Big| \le ||x - y|| \tag{2.1}$$

<u>Dim</u>. Va provato che:

$$-||x - y|| \le ||x|| - ||y|| \le ||x - y|| \,. \tag{2.2}$$

Usando la disuguaglianza triangolare si vede che

$$||x|| \le ||x - y|| + ||y|| \qquad \text{ed anche} \quad ||y|| \le ||y - x|| + ||x|| \,,$$

ossia la (2.2). ∎

E' conseguenza della (2.1):

Teorema 18 *La norma, come trasformazione dallo s.l.n. X a $\mathbb{R}$, è uniformemente continua.*

Inoltre:

Teorema 19 *Sia $f(x)$ una funzione da uno s.l.n. X ad uno s.l.n. Y. Si equivalgono le condizioni*

$$\lim_{x \to x_0} f(x) = 0 \qquad e \qquad \lim_{x \to x_0} ||f(x)||_Y = 0$$

Introduciamo ora la seguente definizione:

Definizione 20 Una successione (x_n) di elementi di X è FONDAMENTALE quando per ogni $\epsilon > 0$ esiste un indice N_ϵ tale che

$$n\,,\ m > N_\epsilon \Longrightarrow ||x_n - x_m|| < \epsilon\,. \ \blacksquare$$

Con la stesa dimostrazione che si conosce per le successioni di numeri reali, si prova che *ogni successione convergente è fondamentale.*

Lo s.l.n. X si dice COMPLETO se ogni successione fondamentale è convergente. ∎

Si vede facilmente che tutte le regole di calcolo dei limiti che valgono in $\mathbb{R}^n$ o in $\mathbb{C}^n$ valgono in qualunque s.l.n. (ma ovviamente il teorema della funzione monotona, che dipende dalla relazione di ordine, non ha corrispondente).

Si sa che sia $\mathbb{R}^n$ che $\mathbb{C}^n$ sono spazi completi per ogni n. Invece:

Teorema 21 *Esistono s.l.n-ti non completi.*

Questo teorema suggerisce:

Definizione 22 Uno spazio lineare *normato e completo* si chiama SPAZIO DI BANACH. ∎

Vedremo in seguito esempi di spazi di Banach. Dovrebbe essere noto, come conseguenza del *teorema dei limiti,* che *lo spazio lineare* $C(a, b)$ *è completo.* Per una dimostrazione si veda il paragrafo 2.3.2.1.

Si sa che esistono successioni (q_n) la cui immagine è densa in $\mathbb{R}^n$ o in $\mathbb{C}^n$. Invece:

Teorema 23 *Esistono s.l.n-ti nei quali nessuna successione ha immagine densa.*

L'ultimo teorema suggerisce la seguente definizione:

Definizione 24 Sia X uno s.l.n. Se esiste una successione (x_n) a valori in X la cui immagine è densa in X, lo spazio X si dice SEPARABILE. ∎

Il teorema 23 può quindi enunciarsi dicendo che *esistono* s.l.n-ti *non separabili.*

Per concludere quest'introduzione, ricordiamo che sul medesimo s.l.n. X possono introdursi più norme. Per esempio su $\mathbb{C}^n$ sono norme tra loro diverse le seguenti. Se $x = \mathrm{col}\begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}$:

$$||x||_p = \left[\sum_{i=1}^{n} |x_i|^p \right]^{1/p} , \quad (\text{se } p \geq 1) \qquad ||x||_\infty = \max\{|x_i| \,,\ i = 1 \dots, n\} \,.$$

(E' facile provare che $|| \cdot ||_1$ e $|| \cdot ||_\infty$ sono norme. Gli altri casi sono più difficili. Si noti però che $|| \cdot ||_2$ è l'usuale norma euclidea).

Tuttavia, su $\mathbb{C}^n$, *la proprietà*

$$\lim x_n = x_0 \tag{2.3}$$

è indipendente dalla particolare norma che si usa per verificarla.

Invece:

Teorema 25 *Esistono s.l.n-ti* X *sui quali si possono definire due diverse norme,* $|| \cdot ||_1$ *e* $|| \cdot ||_2$*, tali che la (2.3) valga per una norma, ma non per l'altra.*

Questo teorema suggerisce di definire:

Definizione 26 Siano $||\cdot||_1$ e $||\cdot||_2$ due norme sul medesimo s.l.n. X. Si dice che le due norme sono EQUIVALENTI se la condizione

$$\lim ||x_n - x_0||_1 = 0 \qquad \text{per ogni } x_0$$

implica la condizione

$$\lim ||x_n - x_0||_2 = 0 \qquad \text{per ogni } x_0,$$

e viceversa. ∎

Osservazione 27 Dunque, due norme equivalenti su X subordinano la stessa topologia su X. ∎

E' importante poter decidere se due norme sono equivalenti. Il teorema seguente dà un test utile:

Teorema 28 *Sia X uno s.l.n. e siano $||\cdot||_1$ e $||\cdot||_2$ due norme su X. Esse sono equivalenti se e solo se esistono due numeri m ed M tali che*

$$m > 0\,, \quad \text{e inoltre} \quad m||x||_1 \leq ||x||_2 \leq M||x||_1\,. \tag{2.4}$$

In particolare, una successione (x_n) che ammette limite x rispetto ad una delle due norme ammette limite anche rispetto all'altra, e questo è ancora x.

Abbiamo notato che esistono spazi lineari *normati e non completi.* E' bene sapere:

Teorema 29 *Sia X uno s.l.n. Si costruisce uno s.l.n. $\hat{X}$ con queste proprietà:*

- *$\hat{X}$ è completo;*
- *un s.spazio X_0 di $\hat{X}$,* denso *in $\hat{X}$, è* isometricamente isomorfo *ad X.*

L'ultima affermazione vuol dire che esiste un isomorfismo J tra X ed X_0 tale che

$$x \in X\,, x = J\tilde{x}\,,\ \tilde{x} \in X_0 \subseteq \hat{X} \Longrightarrow ||x||_X = \|J\tilde{x}\|_X == ||\tilde{x}||_{\hat{X}}\,.$$

2.1.1 Dimostrazioni posposte

Passiamo ora alla dimostrazione dei teoremi che abbiamo enunciato.

Dimostrazione del TEOREMA 21. Un esempio di spazio lineare normato non completo si costruisce come segue: i suoi elementi sono le funzioni $x(\cdot)$ continue su $[-1, 1]$ con l'usuale struttura lineare. La norma è

$$||x(\cdot)|| = \int_{-1}^{1} |x(s)| \, \mathrm{d}s \, .$$

E' facile vedere che la successione (x_n):

$$x_n(t) = \begin{cases} -1 & \text{se} \quad -1 \le t \le -\frac{1}{n} \\ nt & \text{se} \quad -\frac{1}{n} \le t \le \frac{1}{n} \\ 1 & \text{se} \quad \frac{1}{n} \le t \le 1 \end{cases}$$

è fondamentale. Per ogni t, la successione numerica $(x_n(t))$ converge al numero $\mathrm{sgn}(t)$. Essendo la successione (x_n) limitata, segue che

$$\lim \int_{-1}^{1} |x_n(s) - \mathrm{sgn}(s)| \, \mathrm{d}s = 0 \, .$$

Se fosse anche $\lim x_n(\cdot) = \phi(\cdot)$ *nello spazio in cui stiamo lavorando, ossia con* $\phi(\cdot)$ *continua* avremmo

$$\int_{-1}^{1} |\phi(s) - \mathrm{sgn}(s)| \, \mathrm{d}s \le \lim ||\phi - x_n|| + \lim \int_{-1}^{1} |x_n(s) - \mathrm{sgn}(s)| \, \mathrm{d}s = 0 \, ;$$

ossia, $\phi(t) = \mathrm{sgn}(t)$ q.o. $t \in [-1, 1]$. Ciò non può aversi perché $\phi(\cdot)$ dovrebbe essere continua mentre la funzione segno ha un salto. ∎

Dimostrazione del TEOREMA 23. Un esempio di s.l.n. non separabile è il seguente, che si indica col simbolo l^∞: i suoi elementi sono le successioni (x_n) limitate e la norma è

$$||(x_n)||_\infty = \sup |x_n| \, .$$

Sia S una *successione di elementi di* l^∞, $S = (x^{(n)})$. Notiamo che per ogni n il simbolo $x^{(n)}$ indica un elemento di l^∞, ossia una successione $(x_k^{(n)})$ di indice k. Mostriamo che l'immagine di S non è densa in l^∞ costruendo un elemento $y = (y_k) \in l^\infty$ che dista almeno 1 da ciascun elemento $x^{(n)}$ della successione S. Costruiamo y specificandone gli elementi y_k. Per scegliere y_1, primo elemento della successione y, si guarda il primo elemento della successione $x^{(1)}$ e si pone:

$$y_1 = \begin{cases} 0 & \text{se} \quad |x_1^{(1)}| > 1 \\ 2 & \quad\text{altrimenti} \, . \end{cases}$$

In questo modo, qualunque siano i successivi elementi di y, si ha

$$||y - x^{(1)}|| \geq 1\,.$$

Per scegliere y_k si guarda la k-ma successione $x^{(k)}$ e si pone

$$y_k = \begin{cases} 0 & \text{se} \quad |x_k^{(k)}| > 1 \\ 2 & \quad\quad \text{altrimenti}\,. \end{cases}$$

Indipendentemente dai valori degli y_r con $r \neq k$, $||y - x^{(k)}|| \geq 1$. Dunque, la successione $y \in l^\infty$ che abbiamo costruita verifica

$$||y - x^{(k)}|| \geq 1$$

per ogni k; e quindi la successione S non è densa in l^∞. ∎

Si veda il Teorema 37 per un secondo esempio di s.l.n. non separabile.

Dimostrazione del TEOREMA 25. Si consideri lo spazio lineare che indichiamo con X, i cui elementi sono funzioni continue su $[-1,1]$, con le due norme

$$||x||_1 = \int_{-1}^{1} |x(s)|\,\mathrm{d}s\,, \qquad ||x||_\infty = \max_{[-1,1]} |x(t)|\,.$$

La norma $||\cdot||_\infty$ corrisponde alla CONVERGENZA UNIFORME.

Sia (x_n) una successione in X. Se essa converge ad x_0 per $||\cdot||_\infty$, ossia se converge uniformemente, allora vale anche

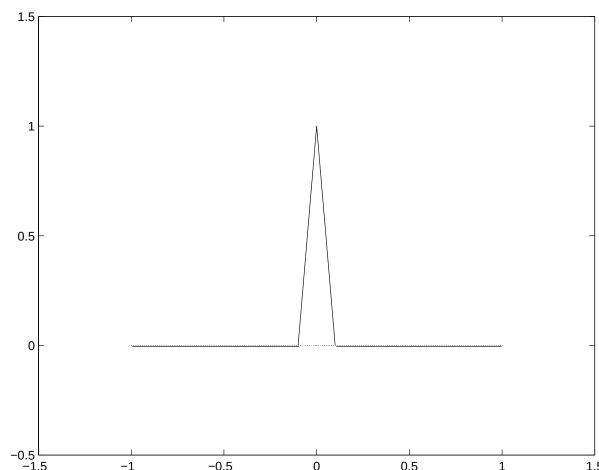
$$\lim \int_{-1}^{1} |x_n(t) - x_0(t)|\,\mathrm{d}t = 0$$

e quindi anche $\lim ||x_n - x_0||_1 = 0$. Esitono però successioni convergenti in $||\cdot||_1$ ma che non sono convergenti nella norma della convergenza uniforme. Sia $x_n(\cdot)$ la funzione il cui grafico è disegnato in figura 2.1:

$$x_n(t) = \begin{cases} 0 & t < -\frac{1}{n} \\ n\left(t + \frac{1}{n}\right) & -\frac{1}{n} \leq t \leq 0 \\ -n\left(t - \frac{1}{n}\right) & 0 \leq t \leq \frac{1}{n} \\ 0 & t > \frac{1}{n}\,. \end{cases}$$

L'area del triangolo tende a zero, e quindi $\lim x_n = 0$ in $||\cdot||_1$; ma la successione (x_n) non è fondamentale in $||\cdot||_\infty$.

Figura 2.1:



<u>*Dimostrazione del TEOREMA 28*</u>. Proviamo la condizione sufficiente. Si sappia che $x_n \to x_0$ in $||\cdot||_2$, ossia si sappia che

$$||x_n - x_0||_2 \to 0\,.$$

Dalla prima disuguaglianza in (2.4) segue

$$m||x_n - x_0||_1 \le ||x_n - x_0||_2 \to 0\,,$$

ossia $x_n \to x_0$ in $||\cdot||_1$ perché $m > 0$.

Se si sa che $x_n \to x_0$ in $||\cdot||_1$, segue la convergenza in $||\cdot||_2$ dalla seconda delle (2.4).

Proviamo il viceversa. Si sappia che *ogni* successione convergente in $||\cdot||_2$ converge, ed ha il medesimo limite, anche in $||\cdot||_1$; ma supponiamo per assurdo che la prima disuguaglianza in (2.4) non valga per nessuna scelta di $m > 0$. In tal caso, per ogni n esiste x_n tale che

$$||x_n||_2 = 1 \qquad \text{e} \qquad ||x_n||_1 > n\,.$$

Definiamo $y_n = x_n/n$. E' ovvio che (y_n) tenda a zero in $||\cdot||_2$ mentre invece $||y_n||_1 \ge 1$ per ogni n; e quindi (y_n) non converge a zero rispetto a $||\cdot||_1$. Questo non si dà, per ipotesi, e quindi esiste $m > 0$ per cui vale la prima diseguaglianza in (2.4).

Analogamente si vede che se la convergenza in $||\cdot||_1$ implica la convergenza in $||\cdot||_2$ allora vale la seconda diseguaglianza in (2.4) per un certo M. ∎

Dimostrazione del TEOREMA 29. La costruzione di $\hat{X}$ è del tutto simile alla costruzione di Cantor dei numeri reali, e viene solamente accennata. Si considera l'insieme delle successioni fondamentali in X e in questo insieme si stabilisce la relazione di equivalenza

$$(x_n) \sim (y_n) \qquad \text{se} \qquad \lim(x_n - y_n) = 0\,.$$

Si vede facilmente che questa è una relazione di equivalenza. Indichiamo con $[(x_n)]$ la classe di equivalenza cui (x_n) appartiene. Si vede facilmente che l'insieme delle classi di equivalenza diviene uno spazio lineare se dotato delle operazioni

$$[(x_n)] + [(y_n)] = [(x_n + y_n)]\,, \qquad \alpha[(x_n)] = [(\alpha x_n)]\,.$$

In questo spazio, che indichiamo con $\hat{X}$, si introduce la norma

$$||\,[(x_n)]\,|| = \lim ||x_n||\,.$$

Questa definizione ha senso perché se (x_n) è fondamentale allora $(||x_n||)$ è una successione fondamentale di numeri, come si vede facilmente dal Teorema 17. Si vede inoltre che la norma così definita dipende dalla classe di equivalenza e non dal rappresentante, ed ha effettivamente le proprietà di una norma su $\hat{X}$.

Lo spazio X_0 è quelo delle classi di equivalenza che hanno un rappresentante costante. E' immediato verificare che X_0 ed X sono isometrici.

Notiamo ora che X_0 è denso in $\hat{X}$ perché ogni elemento $[(x_n)]$ di $\hat{X}$ si approssima mediante la successione $(\,[(y_n)]_r\,)$ così costruita: $y_n = x_n$ se $n < r$; altrimenti $y_n = x_r$. ∎

2.2 Spazi prodotto

Ricordiamo che il PRODOTTO CARTESIANO $X \times Y$ di due insiemi X ed Y è l'insieme delle COPPIE ORDINATE (x, y), il cui primo elemento è in X ed il secondo in Y. Come casi particolari, si possono considerare i casi

$$Y = X \qquad \text{oppure} \qquad X = \Phi\,.$$

Se X ed Y sono spazi lineari, si può rendere $X \times Y$ uno spazio lineare definendo

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)\,, \qquad \alpha(x, y) = (\alpha x, \alpha y)\,.$$

Se inoltre X ed Y sono dotati di norma, rispettivamente $||\cdot||_X$ e $||\cdot||_Y$, si possono definire norme su $X \times Y$. Per esempio

$$||(x,y)||_p = [||x||_X^p + ||y||_Y^p]^{1/p} \quad \forall p \geq 1\,, \qquad ||(x,y)||_\infty = \max\{||x||_X + ||y||_Y\}\,. \tag{2.5}$$

E' un fatto che non proviamo, il seguente: *le norme precedenti su $X \times Y$ sono tutte tra di loro equivalenti.*

Proviamo ora:

Teorema 30 *Sia $X = Y$ e dotiamo $X \times X$ di una qualsiasi delle norme in (2.5). Le due trasformazioni*

$$\begin{array}{ll}(x,y) \to x+y & \textit{da} \quad X \times X \to X \\ (\alpha, x) \to \alpha x & \textit{da} \quad \Phi \times X \to X\end{array}$$

sono continue.

Dim. Dato che le norme in (2.5) sono tutte tra loro equivalenti, la dimostrazione può farsi scegliendo quella che semplifica i calcoli. Per esempio, dotiamo $X \times X$ della norma $\|\cdot\|_1$.

Per provare la prima affermazione, si fissa (x_0, y_0) e si nota che, lavorando con $||\cdot||_1$ su $X \times X$, la disuguaglianza triangolare implica

$$\begin{aligned}&||(x+y) - (x_0+y_0)||_X = ||(x-x_0) + (y-y_0)||_X \leq ||x-x_0||_X + ||y-y_0||_X \\ &= ||(x-x_0, y-y_0)||_1 = ||(x,y) - (x_0,y_0)||_1\,.\end{aligned} \tag{2.6}$$

Segue da qui la continuità della trasformazione $(x,y) \to x+y$.

Per provare la seconda affermazione, fissiamo (α_0, x_0) con $\alpha_0 \neq 0$ e valutiamo:

$$\begin{aligned}&||\alpha x - \alpha_0 x_0||_X = ||(\alpha - \alpha_0)x + \alpha_0(x - x_0)||_X \\ &\leq |\alpha - \alpha_0| \cdot ||x||_X + |\alpha_0| \cdot ||x - x_0||_X\,.\end{aligned}$$

Si ha quindi

$$||\alpha x - \alpha_0 x_0||_X < \epsilon$$

se si sceglie

$$||x - x_0||_X \leq \epsilon/2\alpha_0$$

(e quindi anche $||x||_X \leq ||x_0||_X + \epsilon/2\alpha_0$) ed anche

$$|\alpha - \alpha_0| < \frac{\epsilon}{2[||x_0||_X + \epsilon/2\alpha_0]}\,.$$

In particolare quindi se si sceglie la coppia (α, x) in un intorno di (α_0, x_0) la cui esplicita determinazione si lascia al lettore.

Il caso $\alpha_0 = 0$ si lascia per esercizio. ∎

Osservazione 31 Si noti la stretta somiglianza della dimostrazione del teorema precedente con le usuali dimostrazioni sui limiti della somma e del prodotto di funzioni di una variabile; e si noti, da (2.6), che la somma è anche *uniformemente continua;* una proprietà che invece non vale per il prodotto. ∎

Il teorema precedente implica che in particolare sono continue le trasformazioni

$$x \longrightarrow x_0 + x\,, \qquad x \longrightarrow \alpha_0 x$$

(con x_0 ed α_0 fissati) dette rispettivamente TRASLAZIONI ed OMOTETIE.

2.3 Gli esempi principali di spazi di Banach

Mostriamo ora gli esempi di spazi lineari normati che sono più importanti per le applicazioni. Essi sono tutti spazi completi, ossia spazi di Banach. La dimostrazione della completezza viene vista successivamente.

Sottolineiamo da subito che gli spazi che si incontrano nelle applicazioni hanno un simbolo standard, che indica sia lo spazio vettoriale che la norma su esso.

Prima di elencare gli esempi più importanti di spazi di Banach, richiamiamo le disuguaglianze seguenti, che verranno frequentemente usate in futuro.

DISUGUAGLIANZA DI HÖLDER, rispettivamente per le serie e gli integrali.

Sia $p > 1$. Si definisce[1] ESPONENTE CONIUGATO di p il numero

$$q = \frac{p}{p-1}\,.$$

Siano (a_m) (b_m) successioni, rispettivamente siano $f(x)$ e $g(x)$ funzioni misurabili su Ω.

Supponiamo che sia, rispettivamente,

$$\begin{array}{ll} \sum_{n=0}^{+\infty} |a_n|^p < +\infty\,, & \sum_{n=0}^{+\infty} |b_n|^q < +\infty\,; \\ \int_\Omega |f(x)|^p\,\mathrm{d}x < +\infty\,, & \int_\Omega |g(x)|^q\,\mathrm{d}x < +\infty\,. \end{array} \tag{2.7}$$

Vale rispettivamente:

$$\sum_{n=0}^{+\infty} |a_n b_n| \le \left[\sum_{n=0}^{+\infty} |a_n|^p\right]^{1/p} \left[\sum_{n=0}^{+\infty} |b_n|^q\right]^{1/q}\,,$$

$$\int_\Omega |f(x)g(x)|\,\mathrm{d}x \le \left[\int_\Omega |f(x)|^p\,\mathrm{d}x\right]^{1/p} \left[\int_\Omega |g(x)|^q\,\mathrm{d}x\right]^{1/q}$$

[1] in seguito si estenderà la definizione, includendo anche $p = 1$.

(in queste disuguaglianze il simbolo $0 \cdot \infty$ non compare grazie all'ipotesi (2.7).

Si noti che l'esponente coniugato di $p = 2$ è $q = 2$. La disuguaglianza di Hölder con $p = 2$ si chiama DISUGUAGLIANZA DI SCHWARZ.

DISUGUAGLIANZA DI MINKOWSKI, rispettivamente per le serie e gli integrali. Sia $p \geq 1$ e siano (a_m) (b_m) successioni, rispettivamente siano $f(x)$ e $g(x)$ funzioni misurabili su Ω. Vale

$$\left\{ \sum_{n=0}^{+\infty} |a_n + b_n|^p \right\}^{1/p} \leq \left\{ \sum_{n=0}^{+\infty} |a_n|^p \right\}^{1/p} + \left\{ \sum_{n=0}^{+\infty} |b_n|^p \right\}^{1/p} ,$$

$$\left\{ \int_\Omega |f(x) + g(x)|^p \, \mathrm{d}x \right\}^{1/p} \leq \left\{ \int_\Omega |f(x)|^p \, \mathrm{d}x \right\}^{1/p} + \left\{ \int_\Omega |g(x)|^p \, \mathrm{d}x \right\}^{1/p} .$$

Si noti una conseguenza importante della disuguaglianza di Minkowski: *se ambedue le serie o gli integrali a destra sono finiti anche il membro sinistro è finito.*

Osservazione 32 Sia la disuguaglianza di Hölder che la disuguaglianza di Minkowski si estendono al caso $p = 1$ (l'esponente coniugato è $q = +\infty$) e al caso $p = +\infty$ (l'esponente coniugato è $q = 1$).

Nel caso delle serie le disuguaglianze di Hölder e di Minkowski con $p = 1$ sono:

$$\sum_{n=1}^{+\infty} |a_n b_n| \leq (\sup\{|b_n|\}) \sum_{n=1}^{+\infty} |a_n| ,$$

$$\sum_{n=1}^{+\infty} |a_n + b_n| \leq \sum_{n=1}^{+\infty} |a_n| + \sum_{n=1}^{+\infty} |b_n|$$

(per la disuguaglianza di Hölder si suppone che a destra non compaia il simbolo $0 \cdot \infty$).

Le disuguaglianze corrispondenti valgono anche sostituendo le serie con gli integrali, ma sono poco significative perché l'integrale non dipende dal valore di una funzione in un singolo punto, valore che invece altera in modo significativo l'estremo superiore. L'estensione significativa di queste disuguaglianze al caso degli integrali si vedrà dopo che avremo presentato la definizione di *estremo superiore essenziale.* ∎

2.3.1 Gli esempi di spazi lineari normati

Introduciamo ora gli esempi più importanti di s.l.n. con i simboli comunemente usati per indicarli.

- Il simbolo $C([a,b])$ (più semplicemente $C(a,b)$).

Questo simbolo indica lo s.l.n. i cui elementi sono le *funzioni continue sull'intervallo* $[a,b]$ *chiuso e limitato.* La struttura lineare è quella usuale e la norma è quella della *convergenza uniforme:*

$$||x|| = \max_{[a,b]} |x(t)| \,.$$

In generale, se K è un compatto di Φ^n, con $C(K)$ si intende lo spazio delle funzioni continue su K, con norma $\max_K |x(t)|$.

I valori assunti dalla funzione sono numeri o vettori; talvolta sono matrici.

Osservazione 33 Sottolineiamo nuovamente che ciascuno dei simboli seguenti indica uno spazio lineare con la norma che è ad esso associata nella definizione corrispondente. Quindi, per esempio, non useremo il simbolo $C(a,b)$ per indicare lo spazio delle funzioni continue con una norma integrale, per esempio quella introdotta nella dimostrazione del Teorema 25. Per questo, in quella dimostrazione abbiamo indicato genericamente con X tale s.l.n. ∎

- I simboli l^p, $1 \le p \le +\infty$.

Questi simboli indicano spazi di successioni. Gli elementi sono le successioni (x_n) tali che:

$$\begin{cases} \sum |x_n|^p < +\infty & \text{se } 1 \le p < +\infty, \quad ||(x_n)||_p = \left[\sum |x_n|^p \right]^{1/p} \\ \sup |x_n| < +\infty & \text{se } p = +\infty, \quad ||(x_n)||_\infty = \sup |x_n| \,. \end{cases}$$

E' immediato verificare che gli spazi di successioni appena descritti sono s.l.n-ti rispetto alle usuali operazioni di somma elemento per elemento e di prodotto $\alpha(x_n) = (\alpha x_n)$. La verifica è diretta se $p = 1$ oppure $p = +\infty$ mentre fa uso della disuguaglianza di Minkovski per le serie se $1 < p < +\infty$.

- Il simbolo c_0.

Si usa per per indicare lo s.l.n. delle successioni (x_n) tali che $\lim x_n = 0$. La norma in c_0 è

$$||(x_n)|| = \sup |x_n| \,.$$

Dunque c_0 è un s.spazio di l^∞.

E' facile verificare che sia c_0 che gli spazi l^p con $p < +\infty$ sono *separabili* mentre abbiamo visto, nel corso della dimostrazione del Teorema 23, che l^∞ *non è separabile.*

- I simboli $\mathcal{L}^p(a, b)$ ed $L^p(a, b)$, $1 \le p \le +\infty$.

Il simbolo $\mathcal{L}^p(a, b)$ si usa per indicare gli spazi lineari (non normati!) i cui elementi sono le funzioni $f(\cdot)$ tali che, rispettivamente,[2]

$$\begin{array}{lll} \int_a^b |f(x)|^p \, \mathrm{d}x < +\infty & \text{se} & 1 \le p < +\infty \,, \\ \sup |f(x)| < +\infty & \text{se} & p = +\infty \,. \end{array}$$

E' immediato vedere che $\mathcal{L}^p(a, b)$ con $p = 1$ e $p = +\infty$ sono spazi lineari[3]. Anche gli spazi $\mathcal{L}^p(a, b)$ con $1 < p < +\infty$ sono spazi lineari, grazie alla disuguaglianza di Minkowski.

Però le funzioni

$$f(\cdot) \to \left[\int_a^b |f(x)|^p \, \mathrm{d}x \right]^{1/p}$$

non sono norme su questi spazi: una funzione nulla in tutti i punti salvo uno ha integrale nullo. L'estremo superiore è invece una norma ma ben poco significativa se si vuol considerare applicata a classi di funzioni integrabili. Si rimedia a questo problema definendo una relazione di equivalenza, $f \sim g$, con

$$f \sim g \Longleftrightarrow f(t) = g(t) \qquad \text{q.o. } t \in (a, b).$$

E' chiaro che due funzioni equivalenti hanno i medesimi integrali, finiti o meno. Generalmente non hanno il medesimo estremo superiore. Per questa ragione si definisce l'ESTREMO SUPERIORE ESSENZIALE come segue. Ad ogni $r \ge 0$ si associa l'insieme A_r,

$$A_r = \{t \mid f(t) > r\} \,.$$

Per definizione,

$$\operatorname{ess\,sup} |f(t)| = \sup\{r \mid \lambda(A_r) > 0\}$$

[2]gli integrali vanno intesi nel senso di Lebesgue. Si ricordi che una funzione $f(x)$ è integrabile nel senso di Lebesgue se e solo se $|f(x)|$ lo è.

[3]come notazione, si scrive $\mathcal{L}^\infty(a, b)$, rispettivamente $L^\infty(a, b)$, invece di $\mathcal{L}^{+\infty}(a, b)$ o $L^{+\infty}(a, b)$

dove λ denota la misura di Lebesgue.

Chiaramente, se $f \sim g$ allora le due funzioni f e g hanno il medesimo estremo superiore essenziale.

Si definisce una struttura lineare sull'insieme delle *classi di equivalenza* ponendo

$$[f] + [g] = [f + g]\,, \qquad \alpha[f] = [\alpha f]$$

(è facile vedere che questa definizione dipende solo dalle classi di equivalenza e non dai rappresentanti scelti per definire le operazioni). Quindi, per $1 \leq p \leq +\infty$, si definiscono i simboli $L^p(a, b)$ come gli *spazi lineari* delle *classi di equivalenza* di elementi di $\mathcal{L}^p(a, b)$ (stesso esponente p), *dotati delle norme*

$$||\,[f]\,||_p = \left[\int_a^b |f(x)|^p \,\mathrm{d}x\right]^{1/p} \qquad 1 \leq p < +\infty\,, \qquad ||\,[f]\,||_\infty = \operatorname{ess\,sup} |f(t)|\,.$$

Come si è notato, queste norme dipendono solo dalla classe di equivalenza e non dai rappresentanti usati per calcolarle.

Osserviamo nuovamente che gli spazi sopra introdotti sono s.l.n-ti (grazie alla disuguaglianza di Minkowski per gli integrali se $1 < p < +\infty$).

Notiamo infine le due disuguaglianze

$$\int_a^b |f(x)g(x)| \,\mathrm{d}x \leq (\operatorname{ess\,sup} |g|) \int_a^b |f(x)| \,\mathrm{d}x\,,$$
$$\operatorname{ess\,sup} |f + g| \leq \operatorname{ess\,sup} |f| + \operatorname{ess\,sup} |g|\,.$$

Queste disuguaglianze sono le le disuguaglianze di Hölder *e* Minkovski *nel caso degli "esponenti" 1 e $+\infty$.* Queste suggeriscono di definire come esponente coniugato di $p = 1$ l'"esponente" $q = +\infty$.

Definizioni analoghe si danno per funzioni di più variabili. Se queste sono definite su un insieme K il simbolo che si usa è $L^p(K)$.

Osservazione 34 Mentre nella definizione di $C(K)$ l'insieme K deve essere compatto, tale condizione non è richiesta nella definizione di $L^p(K)$. ∎

Bisogna sapere:

Teorema 35 *Esistono successioni $(x_n(\cdot))$ di funzioni misurabili e limitate su un intervallo $[a, b]$ e tali che: 1) la successione di numeri $(x_n(t))$* non converge *per* nessun *valore di t; ma 2) esiste una funzione integrabile x_0 tale che*

$$\lim_n \int_a^b |x_n(t) - x_0(t)|\, dt = 0\,.$$

Dim. Mostriamo un esempio di successione di funzioni con le proprietà del teorema. In questo esempio, $x_0(t) \equiv 0$. Sia $[a, b] = [0, 1]$.

Costruiamo la successione in due passi; e quindi le funzioni verranno a dipendere da due indici n ed r. Ovviamente sarà possibile riscrivere la successione in modo da farla dipendere da un solo indice.

Al passo n dividiamo l'intervallo $[0, 1]$ in n intervalli uguali, I_0^n, $I_1^n, \ldots$, I_n^{n-1}. definiamo quindi le funzioni $x_{n,r}$, $0 \leq r \leq (n-1)$ come segue[4]:

$$x_{n,r}(t) = (-1)^n \begin{cases} (-1)^r & \text{se } t \in I_r^n \\ 0 & \text{altrimenti.} \end{cases}$$

L'integrale di $|x_{n,r}|$ vale $1/n$, e quindi vale la proprietà 2) con $x_0 = 0$; e si vede immediatamente che in ogni punto t infinite funzioni prendono il valore $+1$, infinite altre prendono il valore -1; e quindi vale anche la proprietà 1). ∎

Osservazione 36 Nonostante il teorema 35, nel seguito *confonderemo sempre le funzioni con le loro classi di equivalenza,* e quindi scriveremo f per indicare la classe di equivalenza $[f]$ di cui f è un rappresentante. Ciò è giustificato dal seguente risultato importante, che non proviamo: *Se*

$$\lim \int_a^b |x_n(t) - x_0(t)| \, dt = 0$$

allora esiste una s.successione $(x_{n,k}(t))$ di $(x_n(t))$ che converge ad $x_0(t)$ q.o. su (a, b). ∎

Se una successione di funzioni $(f_n(x))$ converge in $\mathcal{L}^p(\Omega)$, si dice che la successione CONVERGE IN MEDIA DI ORDINE p. Mostriamo ora:

Teorema 37 *Lo spazio $L^\infty(a, b)$ non è separabile.*

Dim. E' sufficiente mostrare un insieme di elementi di $L^\infty(a, b)$ distanti 1 l'uno dall'altro e che non è numerabile.

Fissiamo $s \in (a, b)$ e sia

$$x_s(t) = \begin{cases} 1 & \text{se} \quad t < s \\ 0 & \text{se} \quad t \geq s\,. \end{cases}$$

Se $s \neq s'$ si ha

$$\sup_{t \in (a,b)} |x_s(t) - x_{s'}(t)| = 1$$

---

[4]il fattore $(-1)^n$ ha poca importanza, ma ha un suo ruolo. Per capirlo si guardino i valori $x_{n,r}(0)$.

e la famiglia di questi elementi di $L^\infty(a,b)$ non è numerabile perché l'insieme dei reali non è numerabile. ∎

Invece, vedremo che gli spazi $L^p(a,b)$ con $p<+\infty$ *sono separabili,* si veda il Teorema 49.

- Il simbolo $W^{k,p}(\Omega)$ (SPAZI DI SOBOLEV).

Sia $\Omega\subseteq\mathbb{R}^n$ un aperto (limitato o meno). Si usa il simbolo $W^{1,p}(\Omega)$ per indicare lo spazio delle (classi di equivalenza di) funzioni $u(\cdot)\in L^p(\Omega)$ tali che esistono funzioni (ossia, classi di equivalenza) g_1, g_2,..., g_n in $L^p(\Omega)$, tali che

$$\int_\Omega u(x)\nabla\phi(x)\,\mathrm{d}x=\int_\Omega\left[\begin{array}{ccc} g_1(x) & \dots & g_n(x)\end{array}\right]\phi(x)\,\mathrm{d}x$$

per ogni funzione ϕ di classe C^∞ a supporto compatto in Ω (il simbolo $\mathcal{D}(\Omega)$ indica lo spazio lineare—non normato—delle funzioni di classe $C^\infty(\Omega)$ a supporto compatto in Ω).

La funzione g_i si chiama la i-ma DERIVATA PARZIALE DEBOLE di u.

Lo spazio $W^{1,p}(\Omega)$ si dota della norma

$$||u||=\left[\int_\Omega|u(x)|^p\,\mathrm{d}x+\sum_{j=1}^n\int_\Omega|g_j(x)|^p\,\mathrm{d}x\right]^{1/p}$$

oppure della norma ad essa equivalente

$$||u||_{L^p(\Omega)}+\sum_{j=1}^n||g_j||_{L^p(\Omega)}\,.$$

Il simbolo $W_0^{1,p}(\Omega)$ indica la chiusura in $W^{1,p}(\Omega)$ di $\mathcal{D}(\Omega)$.

Si prova che $W_0^{1,p}(\Omega)=W^{1,p}(\Omega)$ se e solo se $\Omega=\mathbb{R}^n$.

Se $n=1$ allora esiste una sola derivata parziale debole e vale il seguente teorema, che non proviamo:

Teorema 38 *Se $n=1$ ed $\Omega=(a,b)$, ogni funzione $u\in W^{1,p}(a,b)$ è assolutamente continua; e la sua derivata debole coincide con la derivata usuale, calcolata q.o.*

Se $n=1$ ed $\Omega=(a,b)$, una norma equivalente alle precedenti ed un po' più comoda da usare è

$$\left[|u(a)|^p+\int_a^b|u'(x)|^p\,\mathrm{d}x\right]^{1/p}\,.$$

Osservazione 39 Se $n > 1$ l'esistenza di tutte le derivate deboli *non implica la continuità della funzione.* ∎

Le derivate parziali seconde, in senso debole, si definiscono come derivate prime delle derivate prime; e, analogamente, si definiscono le derivate parziali di ogni ordine, in senso debole. Ovviamente, in generale una funzione avrà derivate parziali fino ad un ordine $k \geq 0$ al più (ricordiamo che la derivata di ordine 0 di una funzione è la funzione stessa).

Lo spazio lineare delle funzioni di classe $L^p(\Omega)$ che hanno tutte le derivate parziali fino all'ordine k, in senso debole, risulta essere uno spazio di Banach quando si doti della norma ottenuta sommando le norme $L^p(\Omega)$ di tutte le derivate parziali fino a quelle di ordine k incluso. Questo spazio si indica col simbolo $W^{k,p}(\Omega)$.

Quando $p = 2$, si usa anche il simbolo $H^1(\Omega)$, invece di $W^{1,2}(\Omega)$. Questo simbolo *non va confuso* con quello che si usa per denotare gli spazi di Hardy, che ora definiamo.

- I simboli $H^p(D)$ e $H^p(\Pi_+)$(SPAZI DI HARDY)

Col simbolo D indichiamo il disco $D = \{z \mid |z| < 1\}$ del piano complesso. Col simbolo $H^p(D)$, $1 \leq p < +\infty$, si indica lo spazio lineare i cui elementi sono le funzioni olomorfe $\phi(z)$ tali che

$$\sup_{r<1} \int_0^{2\pi} |\phi(re^{it})|^p \,\mathrm{d}t < +\infty \qquad \text{con} \qquad ||\phi|| = \sup_{r<1} \left[\frac{1}{2\pi} \int_0^{2\pi} |\phi(re^{it})|^p \,\mathrm{d}t \right]^{1/p} .$$

Col simbolo $H^\infty(D)$ si indica lo spazio delle funzioni olomorfe limitate su D e norma $||\phi|| = \sup_D |\phi(z)|$.

Definizioni analoghe si danno sostituendo a D il semipiano Π_+,

$$\Pi_+ = \{z \mid \Re e\, z > 0\} .$$

In questo caso l'integrazione sulla circonferenza si sostituisce con l'integrazione sulle rette parallele all'asse immaginario: gli elementi di $H^p(\Pi_+)$ sono le funzioni olomorfe in Π_+ per le quali, se $p < +\infty$, è finita la norma

$$||f||_{H^p} = \sup_{x>0} \left[\int_{-\infty}^{+\infty} |\phi(x+iy)|^p \,\mathrm{d}y \right]^{1/p} .$$

Invece, la norma di $H^\infty(\Pi_+)$ è

$$||\phi||_{H^\infty(\Pi_+)} = \sup_{x>0} |\phi(x+iy)| .$$

2.3.2 Le dimostrazioni della completezza

Ricordiamo che si chiama Spazio di Banach uno s.l.n. che è anche completo e che esistono s.l.n-ti che *non* sono completi, si veda il Teorema 21. Inoltre, al paragrafo 2.3.1 abbiamo presentato numerosi esempi di s.l.n-ti.

Vale:

Teorema 40 *Tutti gli s.l.n-ti presentati al paragrafo 2.3.1 sono completi.*

Per provare questo teorema dovremo esaminare separatamente i vari spazi del par. 2.3.1, fissare l'attenzione su una generica successione (x_n) fondamentale e associarle in qualche modo un elemento x_0 dello spazio, che intuiamo essere il limite della successione. Dobbiamo quindi provare che effettivamente $x_0 = \lim x_n$; ossia dobbiamo provare:

a) la funzione x_0 appartiene a X;

b) la funzione x_0 è limite di (x_n) nella norma di X.

Questo richiederà dimostrazioni diverse a seconda dello spazio che consideriamo.

Prima di studiare le singole dimostrazioni, ricordiamo:

Lemma 41 *Sia X uno s.l.n. e sia (x_n) una successione fondamentale in X. La successione (x_n) è limitata, ossia esiste M tale che*

$$||x_n|| < M \qquad \forall n\,.$$

2.3.2.1 Completezza degli spazi l^∞, c_0, $L^\infty(\Omega)$ e $C(K)$

Si ricordi che l'insieme K del simbolo $C(K)$ è un compatto di Φ^n mentre nessuna condizione si impone all'insieme Ω che compare nel simbolo $L^\infty(\Omega)$; e notiamo che sia l^∞ che c_0 sono spazi di funzioni sull'insieme $\Omega = \mathbb{N}$.

Le dimostrazioni della completezza di questi spazi sono tra loro simili, basate sul teorema del doppio limite, provato in appendice.

Indichiamo con X uno degli spazi che stiamo considerando e sia (x_n) una successione fondamentale. Gli spazi che stiamo considerando sono accomunati da questo: sono spazi di funzioni su un certo insieme (indicheremo con t i suoi elementi) e la norma è definita in modo tale che se (x_n) è fondamentale allora ciascuna delle funzioni $(x_n(t))$ è una successione fondamentale di numeri; e quindi converge. Questa affermazione si interpreta se $X = L^\infty(\Omega)$ in questo modo: gli elementi della successione sono classi di equivalenza $[x_n]$ di funzioni.

Si fissa un rappresentante, che ancora indichiamo x_n, in ciascuna classe. Il limite della successione di numeri $(x_n(t))$ esiste q.o. su Ω. Infatti, si tenga presente questo fatto: se la successione $\{[x_n]\}$ è fondamentale, per ogni $\epsilon > 0$ esiste N_ϵ tale che se $n > N_\epsilon$ ed $m > N_\epsilon$ allora

$$\|[x_n] - [x_m]\|_{L^\infty(\Omega)} < \epsilon \,.$$

Scegliamo un rappresentante $x_n(t)$ di ciascuna delle classi $[x_n]$. Allora, esiste un insieme di misura nulla $I_{n,m}$ tale che

$$|x_n(t) - x_m(t)| < \epsilon \qquad \forall x \in \Omega \setminus I_{n,m} \,.$$

Lunione numerabile di insiemi di misura nulla ha misura nulla e quindi

$$|x_n(t) - x_m(t)| < \epsilon \qquad \forall x \in \Omega \setminus I \qquad \text{ove } I = \cup I_{n,m} \,.$$

Dunque la successione $\{x_n(t)\}$ è fondamentale, e quindi convergente, per ogni $t \in \Omega \setminus I$.

L'asserto precedente non vale per successioni fondamentali in $L^p(\Omega)$ con $p < +\infty$.

Dunque, per ogni valore di t (o q.o. su Ω se $X = L^\infty(\Omega)$) si può definire

$$y(t) = \lim x_n(t) \ :$$

una funzione che si spera appartenga allo spazio che stiamo considerando e che sia limite di (x_n). Proviamo:

a) la funzione y appartiene ad X.

Questo è facile se X non è né $C(K)$ né c_0. Infatti in tal caso basta notare che y è una funzione limitata come limite puntuale di una successione (x_n) che, essendo fondamentale, è limitata: $|x_n(t)| < M$ per ogni t e per ogni n. Inoltre, se $X = L^\infty(\Omega)$, la funzione y può costruirsi a partire da un qualsiasi rappresentante delle classi di funzioni $[x_n]$, e ne è limite puntuale q.o.; e dunque è misurabile ed il suo estremo superiore essenziale è finito.

Se $X = C(K)$ allora la continuità di y seguirà dalla convergenza uniforme della successione di funzioni continue (x_n), che proveremo al punto **b),** grazie al Corollario 44.

Sia $X = c_0$. In questo caso ciascuna delle successioni $x_n = x_n(k)$ tende a zero per $k \to +\infty$, e, come proveremo al punto **b),** la successione stessa converge uniformemente ad $y = (y(k))$. Il *teorema del doppio limite,* teorema 43, mostra che $\lim_k y(k) = 0$, ossia che $y \in c_0$.

Proviamo ora

b) la funzione y è limite della successione (x_n).

La funzione y è stata costruita come *limite puntuale* di $x_n(t)$. Mostriamo che in realtà il limite è *nel senso della norma.* Per questo fissiamo $\epsilon > 0$ ed un numero N_ϵ tale che se n, m sono maggiori di N_ϵ allora valga

$$||x_n - x_m|| < \epsilon \,.$$

Valutiamo ora $|y(t) - x_n(t)|$ per $n > N_\epsilon$ come segue:

$$|y(t)-x_n(t)| \leq |y(t)-x_{n+r}(t)|+|x_{n+r}(t)-x_n(t)| \leq |y(t)-x_{n+r}(t)|+\epsilon \,. \quad (2.8)$$

Questa disuguaglianza vale per ogni t e per ogni $r > 0$. Esiste r (dipendente da t) tale che $|y(t) - x_{n+r}(t)| < \epsilon$. Il numero r esiste perché $y(t)$ (per il valore fissato di t) è limite della successione di numeri $(x_{n+r}(t))$ (l'argomento precedente vale q.o. su Ω se $X = L^\infty(\Omega)$).

Dunque,

$$|y(t) - x_n(t)| \leq \inf_r \{|y(t) - x_{n+r}(t)|\} + \epsilon < 2\epsilon \,.$$

2.3.2.2 Completezza dello spazio $H^\infty(D)$

Se (x_n) è una successione fondamentale in $H^\infty(D)$, essa è anche una successione fondamentale in $L^\infty(D)$ e quindi converge in $L^\infty(D)$ ad una funzione x_0, che è limitata. Per provare la completezza di $H^\infty(D)$ basta mostrare che x_0 è olomorfa. Ciò discende dal Teorema di Weierstrass sulla convergenza, uniforme sui compatti, di successioni di funzioni olomorfe.

Ciò prova che $x_0 \in H^\infty$, $x_0 = \lim x_n$, come si voleva.

Non proveremo la completezza degli spazi $H^p(D)$, proprietà che vale per ogni $p \in [1, +\infty]$.

2.3.2.3 Completezza degli spazi l^p con $1 \leq p < +\infty$

Indichiamo col simbolo $(x_n) = (x_n(k))$ una successione di elementi di l^p. Sia essa fondamentale. Da

$$|x_n(k) - x_m(k)| = \left[\sum_{r=0}^{+\infty} |x_n(r) - x_m(r)|^p\right]^{1/p}$$

segue che per ogni k la successione numerica $(x_n(k))$ è fondamentale e quindi convergente. Ciò induce a definire la successione x_0 con

$$x_0(k) = \lim_n x_n(k) \,.$$

Proviamo ora

a) la successione x_0 appartiene ad l^p.

Notiamo per questo che la successione (x_n) di l^p, essendo fondamentale, è limitata: esiste M *indipendente da* n e tale che

$$\left[\sum_{k=0}^{+\infty} |x_n(k)|^p\right]^{1/p} \leq M\,.$$

Segue che per ogni ν vale

$$\left[\sum_{k=0}^{\nu} |x_0(k)|^p\right]^{1/p} = \lim_n \left\{\left[\sum_{k=0}^{\nu} |x_n(k)|^p\right]^{1/p}\right\} \leq M\,.$$

Passando al limite rispetto a ν, si vede che $x_0 \in l^p$.

b) vale $x_0 = \lim x_n$ in l^p.

Si fissi $\epsilon > 0$ e sia $N = N(\epsilon)$ tale che se n, m superano N allora vale

$$||x_n - x_m|| < \epsilon\,.$$

Per n, m maggiori di N_ϵ e per ogni ν vale

$$\left[\sum_{r=0}^{\nu} |x_n(r) - x_m(r)|^p\right]^{1/p} \leq \left[\sum_{r=0}^{+\infty} |x_n(r) - x_m(r)|^p\right]^{1/p} < \epsilon\,.$$

Tenendo fermi ν ed n, si passi al limite per $m \to +\infty$. Si trova:

$$\left[\sum_{r=0}^{\nu} |x_n(r) - x_0(r)|^p\right]^{1/p} \leq \epsilon$$

e questa disuguaglianza vale per ogni ν. Prendendo l'estremo superiore rispetto a ν si vede che, quando $n > N(\epsilon)$, vale

$$||x_n - x_0||_{l^p} \leq \epsilon\,.$$

Questo volevamo provare.

2.3.2.4 Completezza degli spazi $L^p(\Omega)$, $p < +\infty$

In questa parte conviene distinguere tra gli elementi $[x]$ di $L^p(\Omega)$, ossia le classi di equivalenza, e i loro rappresentanti.

Sia $([x_n])$ una successione fondamentale di $L^p(\Omega)$. Il primo passo per mostrarne la convergenza è di costruire una funzione x_0, la cui classe di equivalenza è candidata ad essere limite di $([x_n])$. Gli esempi precedenti suggeriscono di costruire x_0 calcolando il limite puntuale dei valori di opportuni rappresentanti delle classi $[x_n]$. Questo metodo però non può applicarsi nel caso di $L^p(\Omega)$ perché si sa che una successione di funzioni che converge in media può non convergere in nessun punto, si veda il Teorema 35. Usiamo quindi una strategia diversa.

Ricordiamo una proprietà generale delle successioni fondamentali: *una successione fondamentale che ha una s.successione convergente è essa stessa convergente.* Consideriamo una successione $([x_n])$ di elementi di $L^p(\Omega)$ e per ogni classe fissiamo un rappresentante che indichiamo x_n. In questo modo si trova una successione (x_n) di funzioni definite q.o. su Ω, e tali che per ogni $\epsilon > 0$ esiste $N = N(\epsilon)$ con questa proprietà:

$$n > N_\epsilon \implies \left[\int_\Omega |x_n(s) - x_{n+m}(s)| \, \mathrm{d}s \right]^{1/p} < \epsilon \,. \tag{2.9}$$

Facciamo vedere che una successione di funzioni (x_n) con tale proprietà ammette una *sottosuccessione* convergente q.o., che indicheremo col simbolo (y_k).

Dato che $\{[y_k]\}$ è sottosuccessione della successione fondamentale $\{[x_n]\}$ segue che anche $\{[x_n]\}$ converge.

Ricordando la definizione di N_ϵ in (2.9), la sottosuccessione si costruisce con la regola seguente:

- si fissa $\epsilon = 1$. Si sceglie $N(1) + 1$;
- si fissa $\epsilon = 1/2$. Si sceglie $N(1/2) + 1$;
- in generale, con $\epsilon = 1/2^k$. Si sceglie $N(1/2^k) + 1$.

Si considera quindi la sottosuccessione $(y_k) = (x_{N(k)+1})$.

La successione (y_k) gode della seguente proprietà:

$$||y_r - y_{r+1}||_{L^p(\Omega)} \le \frac{1}{2^r} \,.$$

Proviamo che la successione di funzioni (y_k) converge q.o. E' strumentale a ciò introdurre la serie telescopica

$$\sum_{k=1}^{\infty} z_k \,, \qquad z_k = y_{k+1} - y_k \,.$$

Essendo

$$\sum_{k=1}^{m} z_k = y_{m+1} - y_1 \,,$$

per provare la convergenza della successione, basta provare quella della serie.

La costruzione della successione (y_k) implica la convergenza assoluta della serie. Per vederlo, notiamo prima di tutto che

$$\sum_{k=1}^{\infty} \left[\int_\Omega |z_k(s)|^p \, \mathrm{d}s \right]^{1/p} \leq \sum_{k=1}^{\infty} \frac{1}{2^k} = 1 \,. \tag{2.10}$$

Consideriamo ora la serie dei valori assoluti, $g(s) = \sum_{k=1}^{+\infty} |z_k(s)|$ e le sue somme parziali

$$g_n(s) = \sum_{k=1}^{n} |z_k(s)| \,.$$

La successione delle somme parziali è monotona crescente e quindi esiste

$$g(s) = \lim g_n(s) = \sum_{k=1}^{+\infty} |z_k(s)|$$

e inoltre, dalla disuguaglianza di Minkowski e dalla (2.10),

$$\left[\int_\Omega |g_n(s)|^p \, \mathrm{d}x \right]^{1/p} \leq \sum_{k=1}^{n} \left[\int_\Omega |z_k(s)|^p \, \mathrm{d}x \right]^{1/p} < \sum_{k=1}^{+\infty} \frac{1}{2^k} = 1 \,.$$

Dunque, dal *teorema della convergenza monotona,* $|g(s)|^p$ è integrabile su Ω. Questo implica che la funzione $g(s)$ è finita q.o. su Ω. Dunque, per q.o. $x \in \Omega$, la serie numerica

$$\sum_{k=1}^{+\infty} |z_k(s)|$$

converge; e dunque anche la serie numerica

$$\sum_{k=1}^{+\infty} z_k(s)$$

converge q.o. su Ω. Ciò mostra che la successione $(y_k(s))$ converge q.o. su Ω e permette di definire una funzione

$$x_0(s) = \lim y_k(s) = \sum_{k=1}^{+\infty} z_k(s) + y_1(s) \,.$$

La (2.10) mostra che $x_0 \in L^p(\Omega)$. Inoltre,

$$\left[\int_\Omega \left| x_0(s) - \sum_{k=1}^{n} z_k(s) - y_1(s)\right|^p \mathrm{d}s\right]^{1/p} \leq \sum_{k=n+1}^{+\infty} \left[\int_\Omega |z_k(s)|^p \,\mathrm{d}s\right]^{1/p} \leq \sum_{k=n+1}^{+\infty} \frac{1}{2^k} \to 0\,.$$

Ciò prova che la successione di classi di equivalenza $([y_k])$ converge alla classe di equivalenza $([x_0])$. Ciò è quanto volevamo provare.

Osserviamo che, in particolare, abbiamo anche provato un teorema sulla convergenza in media:

Teorema 42 *Sia $1 \leq p < +\infty$. Se una successione di funzioni (x_n) in $\mathcal{L}^p(\Omega)$ converge in media di ordine p ad x_0 allora esiste una sottosuccessione della (x_n) che converge ad x_0 q.o. su Ω.*

2.3.2.5 Completezza degli spazi di Sobolev

Ricordiamo: sia Ω è un aperto di $\mathbb{R}^n$ e siano $u \in L^p(\Omega)$ una funzione a valori scalari e $v \in L^p(\Omega)$ una funzione a valori vettori n–dimensionali. Si dice che

$$v = \nabla u$$

se per ogni ϕ di classe C^∞ ed a supporto compatto in Ω vale

$$\int_\Omega u(s)\nabla\phi(s)\,\mathrm{d}s = -\int_\Omega v(s)\phi(s)\,\mathrm{d}s\,.$$

In questo caso si dice che $u \in W^{1,p}(\Omega)$ e

$$||u||_{W^{1,p}(\Omega)} = \left[\int_\Omega |u(s)|^p\,\mathrm{d}s + \int_\Omega |v(s)|^p\,\mathrm{d}s\right]^{1/p}\,.$$

Dunque, se (u_n) è fondamentale in $W^{1,p}(\Omega)$, le due successioni (u_n) e $(v_n) = (\nabla u_n)$ sono fondamentali in norma $L^p(\Omega)$; e dunque convergenti,

$$u_n \to u_0\,, \qquad v_n \to v_0\,.$$

Passando al limite si trova quindi

$$\begin{aligned} &-\int_\Omega v_0(s)\phi(s)\,\mathrm{d}s = -\lim\int_\Omega \nabla u_n(s)\phi(s)\,\mathrm{d}s \\ &= \lim\int_\Omega u_n(s)\nabla\phi(s)\,\mathrm{d}s = \int_\Omega u_0(s)\nabla\phi(s)\,\mathrm{d}s \end{aligned}$$

per ogni funzione ϕ di classe $C^\infty(\Omega)$, a supporto compatto. Dunque si ha:

$$\begin{cases} \nabla u_0 = v_0 \\ u_n \to u_0 \quad \text{in } L^p(\Omega) \\ \nabla u_n = v_n \to v_0 \quad \text{in } L^p(\Omega) \end{cases}$$

Dunque $u_0 \in W^{1,p}(\Omega)$ ha per gradiente v_0, ed (u_n) converge ad u_0 in $W^{1,p}(\Omega)$.

Ciò prova la completezza di $W^{1,p}(\Omega)$.

2.3.3 Teorema del doppio limite

Il *Teorema del doppio limite* riguarda una successione di funzioni (x_n), definite su un qualsiasi insieme Ω. Per esempio, per ogni n la x_n può essere una successione $(x_n(k))$, oppure può essere una funzione $x_n(t)$ definita su un intervallo $[a, b]$.

Indichiamo genericamente con t gli elementi di Ω.

Le funzioni prendono valori in uno spazio di Banach che indichiamo con X.

Supponiamo che per ogni n esista[5]

$$\lim_t x_n(t) = L_n \in X\,.$$

In questa scrittura può essere $t \to t_0$, $|t| \to +\infty$ o anche $t \to +\infty$ se $\Omega = \mathbb{R}$. Scriviamo quindi genericamente

$$\lim_{t\to\alpha} x_n(t) = L_n \in X\,. \tag{2.11}$$

Vale allora:

Teorema 43 (del DOPPIO LIMITE) *Per ogni n, esista il limite L_n in (2.11) (e sottolineiamo: $L_n \in X$).*

Se (x_n) converge ad x_0 uniformemente *su Ω allora esiste*

$$\lim_{t\to\alpha} x_0(t) = L_0 \in X$$

e si ha

$$L_0 = \lim_n L_n\,;$$

ossia vale

$$\lim_n \left[\lim_{t\to\alpha} x_n(t)\right] = \lim_{t\to\alpha}\left[\lim_n x_n(t)\right]\,.$$

[5] la precisazione "$L_n \in X$" può sembrare ridondante. In realtà non lo è se accade che $X = \mathbb{R}$. In tal caso la condizione "$L_n \in \mathbb{R}$" esclude che il limite sia $\pm\infty$.

Dim. Proviamo prima di tutto l'esistenza di $\lim_{t\to\alpha} x_0(t) = L_0 \in X$. Un teorema dovuto a Cauchy[6] dà una condizione *necessaria e sufficiente* per l'esistenza del limite in X (ossia, se $X = \mathbb{R}$, per l'esistenza del limite *finito*).

$$\lim_{t\to\alpha} x_0(t) = L_0 \in X$$

se e solo se per ogni $\epsilon > 0$ esiste un intorno[7] I_ϵ di α tale che:

$$t', t'' \in I_\epsilon \quad \Longrightarrow \quad |x_0(t') - x_0(t'')| < \epsilon\,.$$

Useremo la parte sufficiente di questa condizione per provare l'*esistenza* di L_0 e, nel corso della dimostrazione, applicheremo la parte necessaria del teorema a ciascuna delle $x_n(t)$, grazie al fatto che l'esistenza dei limiti in (2.11) è nota per ipotesi.

Per provare l'esistenza di L_0, valutiamo $|x_0(t') - x_0(t'')|$ come segue:

$$|x_0(t') - x_0(t'')| \le |x_0(t') - x_n(t')| + |x_n(t') - x_n(t'')| + |x_n(t'') - x_0(t'')|\,.$$

Fissiamo $\epsilon > 0$.

Usando la convergenza *uniforme,* si scelga n così grande da avere, per ogni t,

$$|x_0(t) - x_n(t)| < \epsilon\,.$$

Non abbiamo ancora scelto né t' né t'' ma, comunque essi saranno scelti, certamente avremo

$$|x_0(t') - x_n(t')| < \epsilon\,, \qquad |x_n(t'') - x_0(t'')| < \epsilon\,.$$

Con questo valore di n ormai fissato, si usi l'esistenza del limite

$$\lim_{t\to\alpha} x_n(t) = L_n \in X\,.$$

Il teorema di Cauchy implica l'esistenza di un intorno I_ϵ di α tale che se $t' \in I_\epsilon$, $t'' \in I_\epsilon$, allora

$$|x_n(t') - x_n(t'')| < \epsilon\,.$$

Si noti che I_ϵ dipende dal particolare valore di n *che è stato fissato* e che dipendente a sua volta dal solo ϵ.

---

[6]analogo a quello che correla le successioni fondamentali e le successioni convergenti.

[7]ovvia interpretazione del termine "intorno" quando per esempio $x \in \mathbb{R}$ e $\alpha = +\infty$. In tal caso l'intorno è una semiretta $(r_\epsilon, +\infty)$.

Ricapitoliamo quanto abbiamo visto finora: per ogni $\epsilon > 0$ esiste un intorno I_ϵ di α tale che per ogni coppia di punti t' e t'' in I_ϵ si ha

$$\begin{aligned} &|x_0(t') - x_0(t'')| \leq \\ &\leq \underbrace{|x_0(t') - x_n(t')|}_{\substack{< \epsilon \text{ per } n \text{ opportunamente} \\ \text{grande, grazie alla} \\ \text{convergenza uniforme}}} \\ &+ \underbrace{|x_n(t') - x_n(t'')|}_{\substack{< \epsilon \text{ se } t' \in I_\epsilon,\ t'' \in I_\epsilon \\ \text{per l'esistenza del limite} \\ \lim_{t\to\alpha} x_n(t) = L_n \in X}} \\ &+ \underbrace{|x_n(t'') - x_0(t'')|}_{\substack{< \epsilon \text{ per } n \text{ opportunamente} \\ \text{grande, grazie alla} \\ \text{convergenza uniforme}}} \\ &< 3\epsilon\,. \end{aligned}$$

Il teorema di Cauchy permette di concludere che esiste

$$\lim_{t\to\alpha} x_0(t) = L_0 \in X\,. \tag{2.12}$$

Proviamo ora che

$$L_0 = \lim_n L_n\,.$$

Valutiamo $|L_0 - L_n|$ come segue:

$$|L_0 - L_n| \leq |L_0 - x_0(t)| + |x_0(t) - x_n(t)| + |x_n(t) - L_n|\,.$$

Sia $\epsilon > 0$ fissato. Grazie alla (2.12), esiste un intorno I_ϵ di α tale che

$$t \in I_\epsilon \quad \Longrightarrow \quad |L_0 - x_0(t)| < \epsilon\,.$$

Grazie alla *convergenza uniforme,* esiste un numero N_ϵ tale che *per ogni t vale*

$$n > N_\epsilon \quad \Longrightarrow \quad |x_0(t) - x_n(t)| < \epsilon\,.$$

Dunque, per $t \in I_\epsilon$ e per ogni $n > N_\epsilon$ vale

$$|L_0 - L_n| < 2\epsilon + |x_n(t) - L_n|\,.$$

Il membro sinistro non dipende da t e quindi *per ogni* $n > N_\epsilon$ si ha anche

$$|L_0 - L_n| \leq 2\epsilon + \lim_{t\to\alpha} |x_n(t) - L_n| = 2\epsilon\,.$$

Ciò è quanto dovevamo provare. ∎

In particolare si ha:

Corollario 44 *Una successione (x_n) di funzioni continue definite su un qualunque insieme Ω ivi converga uniformemente alla funzione x_0. Allora x_0 è continua su Ω.*

2.4 Sottospazi di spazi lineari normati

Sia X uno s.l.n., la cui norma indicheremo col simbolo $||\cdot||$. Sia Y un s.spazio di X. Ricordiamo che questo significa

$$\forall y_1\,,\ y_2 \in Y\,,\quad \forall \alpha\,,\ \beta \in \Phi \Longrightarrow \quad \alpha y_1 + \beta y_2 \in Y\,.$$

In particolare, Y stesso è uno spazio lineare e viene ad essere uno s.l.n. se ad Y si restringe la funzione norma definita su X. In tal caso diremo che Y è un s.spazio dello s.l.n. X e diremo che la norma su Y è quella INDOTTA dalla norma di X. *Notiamo che talvolta potrà essere necessario considerare su Y una norma diversa da quella indotta da X. Ciò va sempre esplicitamente specificato per evitare ambiguità. In caso contrario assumeremo che la norma su Y sia quella indotta dalla norma di X.*

Quando X ha dimensione finita, i sottospazi sono le controimmagini di $\{0\}$ sotto l'azione di trasformazioni lineari; e si sa che:

Teorema 45 *In dimensione finita, tutte le trasformazioni lineari sono continue; e quindi tutti i s.spazi sono chiusi.*

Invece, *se X ha dimensione infinita, esso ammette sia s.spazi chiusi che non chiusi.* Esempi di s.spazi chiusi sono ovviamente $\{0\}$ ed X stesso. Vediamo un esempio di s.spazio non chiuso.

Esempio 46 Si considera lo spazio $C(a, b)$ ed in esso il s.spazio Y dei polinomi. E' chiaro che Y non è chiuso perchè la restrizione ad $[a, b]$ della funzione esponenziale è limite uniforme di polinomi. Infatti, la serie di Taylor

$$\sum_{k=0}^{+\infty} \frac{t^k}{k!} = e^t$$

converge uniformemente sui compatti. ∎

E' importante sapere che non soltanto funzioni "regolari" possono approssimarsi uniformemente con polinomi:

Teorema 47 (di Weierstrass) *Sia f una funzione continua su un intervallo limitato e chiuso $[a,b]$,* a valori reali. *Esiste una successioni $\{p_n\}$ di* polinomi a coefficienti reali *che* converge *ad f,* uniformemente *su $[a,b]$.*

La dimostrazione di questo teorema è importantissima perché permette di introdurre il concetto di "identità approssimata". Ad essa è dedicato il paragrafo 2.4.1.

Come conseguenza del Teorema 47 proveremo:

Teorema 48 *Lo spazio $C(a,b)$ è separabile.*

Come sempre, nella notazione $C(a,b)$ è implicito che si lavori su un intervallo *compatto* $[a,b]$. Invece, nel seguente teorema, ancora conseguenza del Teorema 47, né la limitatezza né la chiusura è richiesta.

Teorema 49 *Sia $1 \leq p < +\infty$ e sia (a,b) un intervallo limitato o meno. Si ha che:*

1. *Se (a,b) è limitato, i polinomi sono densi in $L^p(a,b)$;*

2. *lo spazio $L^p(a,b)$ con $p < +\infty$ è separabile.*

Le dimostrazioni dei teoremi precedenti[8] sono posposte.

Una ulteriore proprietà puramente algebrica degli spazi lineari è la seguente: ogni loro s.spazio ammette complementare. Ricordiamo che un spazio lineare Z è un complementare di un spazio lineare Y (ambedue s.spazi di X) se:

$$Z \cap Y = \{0\}\,, \qquad Z + Y = X\,;$$

equivalentemente, se ogni elemento x di X si rappresenta in modo unico come somma di un elemento di Z e di uno di Y.

Quando il s.spazio Z è complementare del s.spazio Y, la dimensione (finita o meno) di Z si chiama la codimensione di Y.

Quando si lavora con spazi normati, è naturale chiedere se tutti i s.spazi *chiusi* ammettano complementare, anch'esso *chiuso.* In dimensione finita la risposta è affermativa. Invece:

Teorema 50 *Esistono s.l.n-ti X, completi, dotati di s.spazi chiusi i quali sono privi di complementare chiuso.*

[8]che potrebbero essere estese al caso di spazi $C(K)$ ed $L^p(K)$ con $K \subseteq \mathbb{R}^n$ anche con $n > 1$

Particolarmente importanti sono i sottospazi *chiusi* di codimensione 1, e anche i s.insiemi della forma

$$x_0 + H\,,$$

con H sottospazio *chiuso* di codimensione 1. Si chiamano tali s.insiemi gli IPERPIANI di X.

La dimostrazione del Teorema 50 consiste nella esplicita costruzione di un s.spazio chiuso privo di complementare chiuso, per esempio di l^p, con $1 \le p < 2$ oppure con $p > 2$. La costruzione è alquanto macchinosa e viene omessa.

Osservazione 51 E' bene ricordare che gli spazi l^p, $1 \le p \le +\infty$, *sono completi.* Si veda il paragrafo 2.3.2 per la dimostrazione. E' anche bene sapere che *ogni s.spazio chiuso di l^2 ammette complementare.* Si veda il paragrafo 227 per la dimostrazione. ∎

2.4.1 Identità approssimate e dimostrazione del teorema di Weierstrass

Una famiglia di funzioni a valori reali $\{h_\nu(s)\}$ che gode delle proprietà **0–3** seguenti si chiama IDENTITÀ APPROSSIMATA. Le proprietà richieste sono:

0. per ciascun valore di ν, la funzione $s \to h_\nu(s)$ è integrabile su $\mathbb{R}$;

1. $h_\nu(s) \ge 0$ per ogni s;

2. $\int_{-\infty}^{+\infty} h_\nu(s)\,\mathrm{d}s = 1$ per ogni $\nu > 0$;

3. per ogni $\epsilon > 0$ si ha:

$$\lim_{\nu\to 0+} \int_{-\infty}^{-\epsilon} h_\nu(s)\,\mathrm{d}s = 0\,, \qquad \lim_{\nu\to 0+} \int_{+\epsilon}^{+\infty} h_\nu(s)\,\mathrm{d}s = 0\,.$$

La ragione del termine "identità approssimata" è espressa dal teorema seguente, che prova che la famiglia $\{h_\nu\}$ "approssima" l'identità rispetto alla convoluzione. Per una giustificazione più generale si veda il paragrafo 4.3.6.

Teorema 52 *Sia $\{h_\nu(s)\}$ un'identità approssimata e sia $f(x)$ una funzione uniformemente continua e limitata su $\mathbb{R}$. Sia $u(\nu, x)$ la funzione*

$$u(\nu, x) = \int_{-\infty}^{+\infty} h_\nu(x-s) f(s)\,ds = \int_{-\infty}^{+\infty} h_\nu(s) f(x-s)\,ds\,.$$

Vale:

$$\lim_{\nu\to 0+} u(\nu,x) = f(x)\,.$$

Il limite è uniforme su $\mathbb{R}$.

Dim. La proprietà **1.** dice che l'integrale di $h_\nu(s)$ vale 1. Dunque, si può scrivere

$$\begin{aligned}&u(\nu,x)-f(x)=\int_{-\infty}^{+\infty} h_\nu(s)[f(x-s)-f(x)]\,\mathrm{d}s\\ &=\int_{-\infty}^{-\epsilon} h_\nu(s)[f(x-s)-f(x)]\,\mathrm{d}s+\int_{\epsilon}^{+\infty} h_\nu(s)[f(x-s)-f(x)]\,\mathrm{d}s\\ &+\int_{-\epsilon}^{+\epsilon} h_\nu(s)[f(x-s)-f(x)]\,\mathrm{d}s\,.\end{aligned}$$

Si fissa un numero σ e si usa *l'uniforme continuità di f su* $\mathbb{R}$ per scegliere $\epsilon=\epsilon_\sigma$ in modo tale che

$$|f(x-s)-f(x)|<\sigma \qquad s\in(-\epsilon,\epsilon)\,.$$

Le proprietà **1.** e **2.** implicano che l'integrale su $(-\epsilon,\epsilon)$ ha modulo minore di σ per ogni $\nu>0$. Infatti,

$$\begin{aligned}&\left|\int_{-\epsilon}^{\epsilon} h_\nu(s)[f(x-s)-f(x)]\,\mathrm{d}x\right|\le\int_{-\epsilon}^{\epsilon} h_\nu(s)|f(x-s)-f(x)|\,\mathrm{d}x\\ &\le\sigma\int_{-\epsilon}^{\epsilon} h_\nu(s)\,\mathrm{d}s\le\sigma\int_{-\infty}^{+\infty} h_\nu(s)\,\mathrm{d}s=\sigma\,.\end{aligned}$$

Fissato tale numero ϵ_σ si usino le proprietà **1.** e **3.** *e la limitatezza di f su* $\mathbb{R}$ per trovare ν_σ tale che, se $\nu\in(0,\nu_\sigma)$, ciascuno degli integrali rimanenti sia in modulo minore di σ. ∎

Le identità approssimate che si incontrano più spesso in pratica si costruiscono come segue: si assegna una funzione integrabile $h(s)$, non negativa. Dividendola per il suo integrale, si può assumere

$$\int_{-\infty}^{+\infty} h(s)\,\mathrm{d}s=1\,.$$

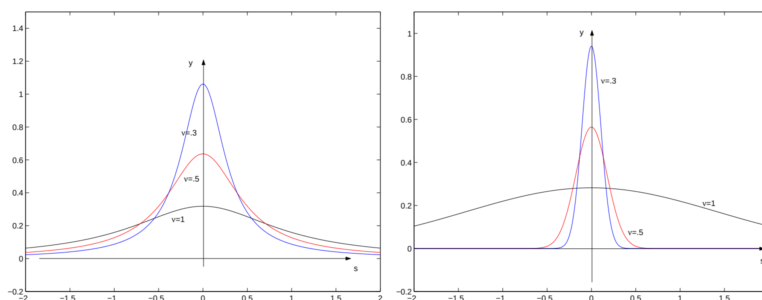
L'identità approssimata si ottiene ponendo

$$h_\nu(s)=\frac{1}{\nu}h(s/\nu)\,. \tag{2.13}$$

La figura 2.2 mostra i grafici di alcune delle funzioni $h_\nu(s)$ così ottenute a partire dalla funzione

$$\frac{1}{\pi(1+s^2)} \quad \text{(sinistra)} \qquad \frac{e^{-s^2}}{\sqrt{\pi}} \quad \text{(destra)}.$$

Figura 2.2: Esempi di identità approssimate



Osservazione 53 Implicitamente abbiamo supposto che ν sia un parametro continuo. Talvolta ν prende valori naturali, e l'identità approssimata è una successione di funzioni. In questo caso il limite per $\nu \to 0+$ si sostituisce col limite per $\nu \to +\infty$.

Diciamo infine che, così come si introducono le identità approssimate su $\mathbb{R}$, si possono anche introdurre le proprietà approssimate su $\mathbb{R}^n$. Le modifiche alla definizione, alla dimostrazione del Teorema 52 (e alla costruzione a partire da una data funzione positiva) sono ovvie e vengono lasciate al lettore. ∎

Veniamo ora alla **dimostrazione del teorema 47,** Teorema di Weierstrass.

La dimostrazione del Teorema di Weierstrass è suggerita da certe considerazioni sull'equazione del calore, che conducono a studiare gli integrali

$$\frac{1}{\sqrt{4\pi t}}\int_{-\infty}^{+\infty} e^{-s^2/4t} f(x-s)\,\mathrm{d}s = u(t,x)\,. \tag{2.14}$$

Come si è notato, la famiglia delle funzioni

$$\frac{1}{\sqrt{4\pi t}} e^{-s^2/4t}$$

è un'identità approssimata (il cui parametro si indica con t perché nelle applicazioni all'equazione del calore indica il tempo. Ovviamente $\sqrt{4t}$ ha ora il ruolo di ν nella (2.13)) ottenuta a partire dalla funzione

$$h(s) = \frac{1}{\sqrt{\pi}} e^{-s^2} \ : \qquad h_t(s) = \frac{1}{\sqrt{4\pi t}} h\left(\frac{s^2}{4t}\right) .$$

Si veda la figura 2.2, a destra.

Sia f una funzione continua su $[a, b]$. Estendiamola in modo qualsiasi ad una funzione continua su $\mathbb{R}$, nulla per $x < a - 1$ e per $x > b + 1$. Indichiamo ancora con f la funzione così estesa.

E' chiaro che la funzione f è uniformemente continua su $\mathbb{R}$. Sia $u(t, x)$ la funzione definita in (2.14). Si vede facilmente che questa funzione è continua per $t > 0$ ed $x \in \mathbb{R}$.

Applichiamo il Teorema 52, ottenendo

$$f(x) = \lim_{t \to 0+} u(t, x) = \lim_{t \to 0+} \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{+\infty} e^{-(x-s)^2/4t} f(s)\, \mathrm{d}s .$$

Ora facciamo intervenire la condizione che f è nulla per $s < a - 1$ e per $s > b + 1$. Si ha così

$$u(t, x) = \frac{1}{\sqrt{4\pi t}} \int_{a-1}^{b+1} e^{-\frac{(x-s)^2}{4t}} f(s)\, \mathrm{d}s .$$

Si fissi $\sigma > 0$ e t_σ tale che

$$|u(t_\sigma, x) - f(x)| < \sigma/2 . \tag{2.15}$$

Il numero t_σ esiste, grazie al Teorema 52.

La disuguaglianza (2.15) vale per ogni $x \in \mathbb{R}$. Limitiamoci però a considerare i valori di x in $[a - 1, b + 1]$. Si rappresenti

$$e^{-\frac{s^2}{4t_\sigma}} = \sum_{k=0}^{+\infty} \frac{1}{k!} \left(-\frac{s^2}{4t_\sigma}\right)^k .$$

La serie, come funzione di s, converge uniformemente su $[a-1, b+1]$. Dunque, esiste N_σ tale che

$$\left| e^{-\frac{s^2}{4t_\sigma}} - \sum_{k=0}^{N_\sigma} \frac{1}{k!} \left(-\frac{s^2}{4t_\sigma}\right)^k \right| < \frac{\sigma\sqrt{4\pi t_\sigma}}{4(b-a)M} , \qquad s \in [a - b - 1, b - a + 1]$$

con $M = \max|f|$.

Si ha quindi:

$$\begin{aligned}
&\left| u(t_\sigma, x) - \frac{1}{\sqrt{4\pi t_\sigma}} \int_{a-1}^{b+1} \sum_{k=0}^{N_\sigma} \frac{1}{k!} \left(-\frac{(x-s)^2}{4t_\sigma} \right)^k f(s)\,\mathrm{d}s \right| \\
&= \frac{1}{\sqrt{4\pi t_\sigma}} \left| \int_{a-1}^{b+1} \left[e^{-\frac{(x-s)^2}{4t_\sigma}} - \sum_{k=0}^{N_\sigma} \frac{1}{k!} \left(-\frac{(x-s)^2}{4t_\sigma} \right)^k \right] f(s)\,\mathrm{d}s \right| \\
&= \frac{1}{\sqrt{4\pi t_\sigma}} \left| \int_{x-b-1}^{x-a+1} \left[e^{-\frac{r^2}{4t_\sigma}} - \sum_{k=0}^{N_\sigma} \frac{1}{k!} \left(-\frac{r^2}{4t_\sigma} \right)^k \right] f(x-r)\,\mathrm{d}r \right| < \sigma/2\,.
\end{aligned}$$

Combinando questa disuguaglianza con (2.15) si vede che la funzione $f(x)$ si approssima uniformemente su $[a, b]$ mediante i polinomi

$$\frac{1}{\sqrt{4\pi t_\sigma}} \int_{a-1}^{b+} \sum_{k=0}^{N_\sigma} \frac{1}{k!} \left(-\frac{(x-s)^2}{4t_\sigma} \right)^k f(s)\,\mathrm{d}s\,. \quad \blacksquare$$

Il teorema si estende a funzioni di n variabili, sostanzialmente con la medesima dimostrazione. Si ricorre per questo al seguente risultato:

Teorema 54 (di Tietze) *Ogni funzione continua su un compatto di $\mathbb{R}^n$ ammette estensione uniformemente continua e limitata ad $\mathbb{R}^n$.*

Si fa quindi uso dell'identità approssimata

$$h_t(x) = \frac{1}{(4\pi t)^{n/2}} e^{\frac{|x|^2}{4t}}\,. \tag{2.16}$$

Si noti che troncando la serie della funzione $e^{|x|^{2k}}$ si trovano polinomi in $|x|^2 = \sum_{i=1}^n x_i^2$.

Osservazione 55 I punti $x \in \mathbb{R}^2$ si possono anche leggere come punti $x + iy$ del piano complesso e gli argomenti precedenti possono adattarsi al caso delle funzioni a valori complessi. Però, gli approssimanti che si ottengono troncando la serie di Taylor dell'identità approssimata (2.16) sono polinomi in $x^2 + y^2$ e *non* in $z = x + iy$. Non si trova quindi un'approssimazione mediante polinomi della variabile complessa z. $\blacksquare$

2.4.2 Dimostrazioni posposte

In questo paragrafo proviamo i Teoremi 48 e 49.

Dimostrazione del TEOREMA 48. Abbiamo provato che l'insieme di *tutti* i polinomi è denso in $C(a,b)$. Quest'insieme non è numerabile. Infatti i polinomi sono parametrizzati dai loro coefficienti, che sono numeri reali. Consideriamo però i soli polinomi che hanno numeri razionali come coefficienti.

Ogni polinomio si approssima in modo uniforme mediante polinomi a coefficienti razionali; e quindi anche l'insieme dei polinomi a coefficienti razionali è denso in $C(a,b)$. Per concludere la dimostrazione, basta notare che l'insieme dei polinomi a coefficienti razionali è numerabile come unione della famiglia numerabile di insiemi $\mathcal{P}_n$, ciascuno numerabile,

$$\mathcal{P}_n = \{\text{polinomi di grado } n \text{ e coefficienti razionali}\} \,. \quad \blacksquare$$

Dimostrazione del TEOREMA 49.

Abbiamo bisogno di alcune proprietà delle funzioni integrabili secondo Lebesgue.

Sia $f \in L^1(a,b)$ e si definisca

$$g(x) = \int_c^x f(s)\,\mathrm{d}s$$

con $c \in [a,b]$. Si prova che la funzione $g(x)$ è continua e q.o. derivabile, con derivata q.o. uguale ad $f(x)$.

Ciò vale sia su intervalli limitati che illimitati.

Supponiamo ora che (a,b) sia un intervallo limitato e che $f(x)$ sia limitata,

$$|f(x)| < M \,.$$

Al variare di h, consideriamo i rapporti incrementali[9]

$$\Phi_h(x) = \frac{g(x+h) - g(x)}{h} = \frac{1}{h}\int_x^{x+h} f(s)\,\mathrm{d}s \,.$$

E' immediato verificare che

$$|\Phi_h(x)| < M$$

[9]Ovviamente interessa il caso $|h|$ piccolo. Per poter definire il rapporto incrementale per ogni $x \in (a,b)$, supponiamo di aver esteso $f(x)$ ad un intervallo un po' più grande, per esempio estendendola con zero.

e, come si è già detto,

$$\lim_{h\to 0} \Phi_h(x) = f(x) \qquad \text{q.o. } x \in (a,b)\,.$$

Dunque,

$$\lim_{h\to 0} \Phi_h(x) = f(x)$$

vale nella norma di $L^p(a,b)$, per ogni $p \in [0,+\infty)$.

Fatta questa premessa, proviamo il teorema 49.

Ricordiamo che (a,b) può non essere limitato e che $f(x)$ può non essere limitata. Premettiamo

Lemma 56 *Sia (a,b) un intervallo limitato e sia $f \in L^p(a,b)$, $1 \le p < +\infty$. Esiste una successione $(p_n(x))$ di polinomi che converge ad $f(x)$ in $L^p(a,b)$.*

Dim. Fissato il numero naturale n, mostriamo come costruire il polinomio $p_n(x)$.

Se la funzione $f(x)$ non è limitata, si determini k in modo tale che $||\tilde{f}(x) - f(x)||_{L^p(a,b)} < 1/3n$, con

$$\tilde{f}(x) = \begin{cases} f(x) & \text{se} \quad |f(x)| < k \\ k & \qquad\text{altrimenti.} \end{cases}$$

Se $f(x)$ è limitata, sia $\tilde{f}(x) = f(x)$.

Introduciamo ora i rapporti incrementali

$$\Phi_h(x) = \frac{1}{h}\int_x^{x+h} \tilde{f}(s)\,\mathrm{d}s\,.$$

Da quanto detto sopra si deduce che per ogni n esiste un valore h_n di h tale che

$$||\tilde{f}(x) - \Phi_{h_n}(x)||_{L^p(a,b)} < \frac{1}{3n}.$$

Per ogni h la funzione $\Phi_{h_n}(x)$ è continua su $[a,b]$ (limitato e chiuso). Per il Teorema di Weierstass, esiste un polinomio $p_n(x)$ tale che

$$\max_{[a,b]} |p_n(x) - \Phi_{h_n}(x)| < \frac{1}{3n(b-a)^{1/p}}$$

e quindi tale che

$$||p_n(x) - \Phi_{h_n}(x)||_{L^p(a,b)} < \frac{1}{3n}.$$

Per questo polinomio si ha

$$||p_n(x) - f(x)||_{L^p(a,b)} \leq ||p_n(x) - \Phi_{h_n}(x)||_{L^p(a,b)}$$
$$+||\Phi_{h_n}(x) - \tilde{f}(x)||_{L^p(a,b)} + ||\tilde{f}(x) - f(x)||_{L^p(a,b)} < \frac{1}{n}\,.$$

Ciò completa la dimostrazione. ∎

Presentiamo ora la dimostrazione del Teorema 49. Procediamo per gradi nella dimostrazione. Introduciamo $\epsilon > 0$ e un numero $R > 0$. Definiamo quindi

$$f_R(x) = \begin{cases} f(x) & \text{se} \quad |x| < R \\ 0 & \quad\quad \text{altrimenti.} \end{cases}$$

Si sa che si può trovare R tale da avere

$$||f - f_R||_{L^p(a,b)} < \epsilon/3\,.$$

Fissato questo valore per R, e questa funzione f_R, basta usare il Lemma 56 per dedurre l'esistenza di un polinomio $p(x)$ tale che

$$||f_R(x) - p(x)||_{L^p(-R,R)} < \epsilon/3\,.$$

La $f(x)$ è quindi approssimata a meno di $(2/3)\epsilon$ dalla funzione

$$\begin{cases} p(x) & \text{se} \quad |x| < R \\ 0 & \quad\quad \text{altrimenti.} \end{cases}$$

Per concludere la dimostrazione è ora sufficiente imporre ad R di prendere valori interi e quindi approssimare i polinomi $p(x)$ mediante polinomi a coefficienti razionali, come nella dimostrazione del Teorema 48.

2.5 La compattezza

Ricordiamo che un s.insieme K dello s.l.n. X si dice RELATIVAMENTE COMPATTO quando *ogni* successione (x_n) a valori in K ammette s.successioni (x_{n_k}) convergenti (non si richiede che il limite appartenga a K). L'insieme K si dice COMPATTO quando è relativamente compatto e chiuso. Il teorema di Bolzano-Weierstrass può riformularsi dicendo che *se* $\Phi = \mathbb{R}$ *oppure* $\Phi = \mathbb{C}$ *allora ogni s.insieme limitato e chiuso di* Φ^n *è compatto.* Si ricordi che questa proprietà è cruciale per la dimostrazione del teorema di Weierstrass sull'esistenza di massimi e minimo.

Sfortunatamente, l'analogo del Teorema di Bolzano-Weierstrass *non* vale in spazi di Banach di dimensione infinita; anzi:

Teorema 57 *Sia X uno spazio normato. Se una sfera*

$$\{x \mid |x - x_0| = \epsilon\}$$

è compatta allora lo spazio ha dimensione finita.

Il Teorema 57 implica in particolare:

Corollario 58 *Sia* $\dim X = +\infty$*. Se K è relativamente compatto, allora K non ha punti interni.*

Infatti, se K contiene punti interni esso contiene una sfera chiusa, che deve essere compatta perché ogni s.insieme chiuso di un compatto è esso stesso compatto. Ciò non può darsi se $\dim X = +\infty$.

E' però vero che:

Teorema 59 *Se il s.insieme K di X è compatto, allora esso è limitato.*

Infatti, se K è illimitato esso contiene l'immagine di una successione (x_n) tale che $||x_n|| \to +\infty$; e si vede facilmente che (x_n) non ammette s.successioni convergenti.

Ricapitolando, in dimensione infinita gli insiemi compatti (rispetto alla topologia della norma) sono pochi (e ciò avrà conseguenze nefaste nei problemi di ottimizzazione) e difficili da caratterizzare. Di conseguenza i pochi teoremi che caratterizzano gli insiemi compatti sono importanti. Il prototipo ed il più utile di essi è il *Teorema di Ascoli-Arzelà,* che caratterizza gli insiemi compatti di $C(a, b)$ (ricordiamo che con questo simbolo si intende in particolare che l'intervallo $[a, b]$ è limitato e chiuso).

Sia K compatto contenuto in $C(a,b)$. Abbiamo già notato, nel Teorema 59 che K deve essere limitato; ossia deve esistere $R>0$ tale che

$$||x|| = \max_{[a,b]} |x(t)| < R \qquad \forall x \in K\,.$$

Trattandosi di limitatezza nella norma della convergenza uniforme, usa anche dire che K è UNIFORMEMENTE LIMITATO.

Ricordiamo che ogni funzione $x \in C(a,b)$ è *uniformemente continua* perché $[a,b]$ è limitato e chiuso; ossia, per ogni $\epsilon > 0$ esiste un numero δ, che dipende da ϵ *e dalla funzione* x, tale che

$$|x(t') - x(t'')| < \epsilon\,, \qquad \forall t' \ \ t'' \in [a,b] \quad \text{per cui } |t'-t''| < \delta.$$

Si dice che l'insieme K è equicontinuo quando δ si può scegliere dipendente da ϵ ma non dall'elemento $x \in K$; ossia:

Definizione 60 Si dice che l'insieme K è EQUICONTINUO quando *per ogni* $\epsilon > 0$ *esiste* $\delta > 0$ tale che *per ogni* $x \in K$ e *per ogni* t', t'' in $[a,b]$ tali che $|t'-t''| < \delta$ si ha:

$$|x(t') - x(t'')| < \epsilon\,. \ ∎$$

Ovviamente, *ogni insieme finito di funzioni continue è equicontinuo.*

Vale:

Teorema 61 (di ASCOLI-ARZELÀ) *Gli insiemi relativamente compatti di* $C(a,b)$ *sono tutti e soli quelli uniformemente limitati ed equicontinui.*

2.5.1 Dimostrazioni posposte

Proviamo i teoremi enunciati.

Dimostrazione del TEOREMA 57. Basta mostrare una successione di elementi di norma 1 priva di sottosuccessioni fondamentali. Notiamo il seguente lemma (di immediata dimostrazione) che verrà utile anche in seguito:

Lemma 62 *Sia* (x_n) *una successione tale che* $||x_n - x_m|| \geq \sigma > 0$ *per ogni coppia* n *ed* m. *La successione* (x_n) *non ha s.successioni fondamentali.*

Per costruire una successione (x_n) con le proprietà richieste dal lemma, facciamo intervenire:

Lemma 63 (di RIESZ) *Sia* M *un s.spazio chiuso di* X, *diverso da* X. *Per ogni* $\epsilon > 0$, *esiste un elemento* $x \in X$ *di norma* 1, *che dista da* M più *di* $1-\epsilon$.

Dim. Si fissi un qualsiasi $\epsilon > 0$ e quindi si fissi un $\sigma > 0$ tale che

$$1 - \epsilon < \frac{1}{1+\sigma}\,. \tag{2.17}$$

Ricordiamo che la distanza $d(x, M)$ di x da M è

$$d(x, M) = \inf\{||x - m|| \mid m \in M\}\,.$$

Essendo $M \neq X$, ed M chiuso, esiste $x_1 \neq M$ a distanza positiva da M:

$$d(x_1, M) = d > 0\,.$$

La definizione di distanza mostra che, essendo $\sigma > 0$, esiste $v_\sigma \in M$ tale che

$$||x_1 - v_\sigma|| < d(1+\sigma)\,.$$

Sia

$$y = x_1 - v_\sigma\,.$$

Ovviamente,

$$||y|| = ||x_1 - v_\sigma|| \le d(1+\sigma) \tag{2.18}$$

e inoltre $d(y, M) = d(x_1, M)$ perché $v_\sigma \in M$; ossia

$$d(y, M) = d(x_1, M) = d \ge \frac{1}{1+\sigma}||x_1 - v_\sigma|| = \frac{1}{1+\sigma}||y||\,.$$

Scegliamo ora

$$x_0 = \frac{y}{||y||}$$

e notiamo che, usando (2.17) e (2.18) si trova:

$$\begin{aligned}
d(x_0, M) &= d\left(\frac{y}{||y||}, M\right) = \inf\left\{\left|\left|\frac{y}{||y||} - m\right|\right| \mid m \in M\right\} \\
&= \inf\left\{\frac{1}{||y||}\,||y - (m||y||)\,|| \mid m \in M\right\} \\
&= \inf\left\{\frac{1}{||y||}\,||x_1 - (v_\sigma + m||y||)\,|| \mid m \in M\right\} \\
&= \inf\left\{\frac{1}{||y||}\,||x_1 - m|| \mid m \in M\right\} = \frac{d}{||y||} > \frac{1}{1+\sigma}\,.
\end{aligned}$$

La dimostrazione è quindi completata. ∎

E' ora facile costruire una successione di elementi di norma 1, priva di s.successioni convergenti: si sceglie $x_1 \neq 0$ qualsiasi e si definisce

$$M_1 = \mathrm{span}\{x_1\} = \{tx_1 \mid t \in \Phi\}\,.$$

Si sceglie quindi x_2, di norma 1, con

$$d(x_2, M_1) > \frac{1}{2}\,.$$

In particolare vale $||x_2 - x_1|| > 1/2$.

Definiti $x_1, \ldots, x_k$, si sceglie x_{k+1} di norma 1, distante almeno 1/2 dallo spazio generato dai vettori $x_1, \ldots, x_k$.

Essendo $\dim X = +\infty$, $X \neq \mathrm{span}\,\{x_1\,,\ \ldots\ \ x_k\}$ e quindi questa costruzione si può ripetere e conduce ad una successione (x_n) i cui elementi distano due a due almeno 1/2, e quindi priva di s.successioni convergenti. ∎

Dimostrazione del TEOREMA 61. Proviamo la parte necessaria. Si è già detto che se K è compatto allora deve essere limitato (teorema 59). Proviamo che se è compatto in $C(a, b)$ allora esso è anche equicontinuo. Si usa il lemma seguente:

Lemma 64 *Sia K compatto in uno spazio di Banach X e sia $\epsilon > 0$. Esiste un insieme* finito *di elementi $k_1, \ldots k_s$ di K, tali che*

$$K \subseteq \bigcup B(k_i, \epsilon)\,, \qquad B(k_i, \epsilon) = \{x \in X \mid ||x - k_i|| < \epsilon\}\,.$$

Dim. per assurdo sia K compatto e sia $\epsilon_0 > 0$ un numero tale che la proprietà non valga. Scelto un qualsiasi $x_1 \in K$, $B(x_1, \epsilon_0)$ non copre K; e quindi esiste $x_2 \in K$ che dista da x_1 più di ϵ_0. Ancora perchè ϵ_0 non soddisfa alla proprietà detta nel lemma, $B(x_1, \epsilon_0) \cup B(x_2, \epsilon_0)$ non copre K. Dunque esiste x_3 in K che dista più di ϵ_0 sia da x_1 che da x_2.

Iterando questo procedimento, si trova una successione (x_n) i cui punti distano l'uno dall'altro almeno ϵ_0, e quindi priva di s.successioni convergenti. Ciò contrasta con la compattezza di K. ∎

Proviamo ora che l'insieme K, compatto in $X = C(a, b)$, è equicontinuo. Si fissi $\epsilon > 0$ e si fissino $k_1, \ldots, k_r$ tali che

$$K \subseteq \bigcup B(k_i, \epsilon)\,.$$

Ciascuna delle funzioni k_i è una funzione uniformemente continua e quindi l'insieme delle k_i, che sono in numero finito, è equicontinuo: esiste $\delta > 0$ tale che se $|t' - t''| < \delta$ allora $|k_i(t') - k_i(t'')| < \epsilon$ per ogni i.

Sia ora $x \in K$ qualsiasi e k_{i_0} una funzione dell'insieme $\{k_1, \dots, k_r\}$ che dista da x meno di ϵ. Valutiamo:

$$|x(t') - x(t'')| \le |x(t') - k_{i_0}(t')| + |k_{i_0}(t') - k_{i_0}(t'')| + |k_{i_0}(t'') - x(t'')| < 3\epsilon .$$

Questa disuguaglianza vale per $|t' - t''| < \delta$ e per ogni $x \in K$. Notando che δ non dipende da $x \in K$ segue l'equicontinuità.

Proviamo ora la condizione sufficiente: proviamo che se $K \subseteq C(a, b)$ è sia limitato che equicontinuo allora K è relativamente compatto. Per questo, scegliamo una qualsiasi successione (x_n) in K e diamo un metodo per costruirne una s.successione convergente. Per fare ciò, fissiamo prima di tutto una successione iniettiva (t_n) la cui immagine è densa in $[a, b]$. Procediamo ora per passi: consideriamo la successione *di numeri* $(x_n(t_1))$. Questa è una successione di numeri *limitata* perché K è limitato. Dunque ammette una s.successione convergente. Indichiamo col simbolo $x_n^{(1)}(t_1)$ questa successione *di numeri* e consideriamo la successione *di funzioni* $(x_n^{(1)})$. Valutiamo queste funzioni nel punto t_2 ottenendo la successione *di numeri* $(x_n^{(1)}(t_2))$. Estraiamo da questa una successione convergente, che indichiamo col simbolo $(x_n^{(2)}(t_2))$. Consideriamo quindi la successione *di funzioni* $(x_n^{(2)})$ e la successione *di numeri* $(x_n^{(2)}(t_3))$. Iteriamo il procedimento.

In questo modo si definiscono induttivamente le successioni *di funzioni* $(x_n^{(1)})$, $(x_n^{(2)})$,..., ciascuna delle quali è s.successione delle precedenti. Dunque, la successione $(x_n^{(k)}(t_r))$ è una successione *di numeri* che converge per ogni $r \le k$. Inoltre, $(x_n^{(1)})$ è s.successione della (x_n).

Consideriamo ora la tabella seguente.

$x_1^{(1)}$	$x_2^{(1)}$	$x_3^{(1)}$	$x_4^{(1)}$	$x_5^{(1)}$	$\dots$
$x_1^{(2)}$	$x_2^{(2)}$	$x_3^{(2)}$	$x_4^{(2)}$	$x_5^{(2)}$	$\dots$
$x_1^{(3)}$	$x_2^{(3)}$	$x_3^{(3)}$	$x_4^{(3)}$	$x_5^{(3)}$	$\dots$
$x_1^{(4)}$	$x_2^{(4)}$	$x_3^{(4)}$	$x_4^{(4)}$	$x_5^{(4)}$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Si ha:

- in ogni casella compare una delle funzioni della successione;
- gli elementi della prima riga costituiscono una s.successione della (x_n);
- ciascuna delle successive righe contiene gli elementi di una s.successione di quella che compare alla riga precedente;

- se si calcolano gli elementi della riga i–ma per $t = t_j$, con $j \le i$, si trova una successione di numeri che converge.

Queste proprietà implicano che la successione diagonale $\left(x_n^{(n)}(\cdot)\right)$ *è una successione* di funzioni *con questa proprietà: le successioni* di numeri $n \mapsto x_n^{(n)}(t_k)$ *convergono, per ogni* k.

Notiamo che per ora abbiamo usato la sola limitatezza dell'insieme K. Usiamo ora l'equicontinuità della successione di funzioni per provare che la successione diagonale $\left(x_n^{(n)}(\cdot)\right)$ è fondamentale (e quindi convergente in $C(a,b)$ che, come si è detto, è uno spazio completo).

Sia $\epsilon > 0$. Si vuol provare l'esistenza di N_ϵ tale che se n, m sono maggiori di N_ϵ allora vale

$$||x_n^{(n)} - x_m^{(m)}|| < 3\epsilon \qquad \text{ossia} \quad |x_n^{(n)}(t) - x_m^{(m)}(t)| < 3\epsilon \quad \forall t \in [a,b].$$

Si fissi $\delta > 0$ tale che se $|t' - t''| < \delta$ allora *ogni* $x \in K$ verifica

$$|x(t') - x(t'')| < \epsilon\,.$$

Rappresentiamo l'intervallo $[a,b]$, che è *limitato,* come unione finita di intervalli di lunghezza minore di δ:

$$[a,b] = \bigcup_{s=1}^{\nu} [a_s, b_s]\,, \qquad b_s - a_s < \delta\,.$$

Per ciascun intervallo $[a_s, b_s]$, fissiamo uno dei punti della successione (t_n) che gli appartiene. Indichiamolo col simbolo $\bar{t}_s$.

Sia $t \in [a,b]$ qualsiasi e sia s tale che $t \in [a_s, b_s]$. Valutiamo, *usando l'equicontinuità,*

$$\begin{aligned}&|x_n^{(n)}(t) - x_m^{(m)}(t)| \le |x_n^{(n)}(t) - x_n^{(n)}(\bar{t}_s)| + |x_n^{(n)}(\bar{t}_s) - x_m^{(m)}(\bar{t}_s)| \\ &+|x_m^{(m)}(\bar{t}_s) - x_m^{(m)}(t)| \le 2\epsilon + |x_n^{(n)}(\bar{t}_s) - x_m^{(m)}(\bar{t}_s)|\,.\end{aligned}$$

Per ogni s, esiste N_s tale che, se n, m sono maggiori di N_s, vale

$$|x_n^{(n)}(\bar{t}_s) - x_m^{(m)}(\bar{t}_s)| < \epsilon$$

ed i punti $\bar{t}_s$ sono in numero finito e non dipendono dal punto t. Dunque, se sia n che m sono maggiori di $N_\epsilon = \max\{N_1, \dots, N_\nu\}$ si ha

$$|x_n^{(n)}(t) - x_m^{(m)}(t)| \le 3\epsilon \qquad \forall t \in [a,b]\,.$$

La dimostrazione è quindi completata. ∎

Il procedimento di estrarre la successione diagonale, dovuto a Cantor, va sotto il nome di METODO DIAGONALE DI CANTOR.

2.6 Operatori lineari

Siano X ed Y due s.l.n-ti e sia f una trasformazione da X in Y. Non si richiede che il dominio di f sia tutto X. Per dire che f opera tra due s.l.n-ti, si dice che f è un OPERATORE.

Si chiamano FUNZIONALI le trasformazioni che operano da X, s.l.n. sul campo scalare Φ, e prendono valori nel campo scalare Φ stesso.

Siano ora X ed Y due spazi lineari *sul medesimo campo scalare.* Si dice che f è un OPERATORE LINEARE da X in Y quando

- $\mathrm{dom}\, f$ è un *sottospazio* (non si richiede chiuso) di X;
- vale $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$ *per ogni* x, y in X e *per ogni* α, β in Φ.

Quando si lavora con operatori lineari, invece della notazione $f(x)$ usa indicare l'operatore con una lettera maiuscola, per esempio F, A; e scrivere Fx, Ax invece di $F(x)$, $A(x)$; ossia si usa la "notazione moltiplicativa" nota dall'algebra lineare.

Studiamo ora le proprietà degli operatori lineari.

2.6.1 Proprietà geometriche degli operatori lineari

Si ricordi che il GRAFICO di una trasformazione $y = f(x)$ da un insieme X ad un insieme Y è l'insieme delle coppie

$$\{(x, y) \mid y = f(x)\} .$$

Invece, l'IMMAGINE ed il NUCLEO[10] sono rispettivamente

$$\mathrm{im}\, f = \{y \mid y = f(x)\} \subseteq Y , \qquad \ker f = \{x \mid f(x) = 0\} \subseteq X .$$

Si provi per esercizio:

Teorema 65 *Siano X ed Y due s.l.n-ti e sia f una trasformazione da X in Y. Vale:*

- *la trasformazione f è lineare se e solo se il suo grafico è un s.spazio di $X \times Y$;*
- *se f è lineare sia la sua immagine che il suo nucleo sono s.spazi.*

[10] il termine "nucleo" per indicare l'insieme degli zeri di una funzione si usa solamente se la funzione è lineare.

Notiamo ora:

Teorema 66 *Un operatore lineare da X in Y che ha immagine limitata è identicamente zero.*

Dim. Infatti, l'immagine di un operatore lineare è un sottospazio: questo è limitato se e solo se è il sottospazio 0. ∎

Col simbolo $B(x_0, r)$ indichiamo la PALLA

$$B(x_0, r) = \{x \mid ||x - x_0|| \le r\} .$$

Lemma 67 *Sia A un operatore lineare da X in Y e sia $x_0 \in \mathrm{dom}A$. Vale*

$$B(x_0, r) \cap (\mathrm{dom}A) = x_0 + B(0, r) \cap (\mathrm{dom}A) .$$

Dim. Infatti, con $L = \mathrm{dom}A$,

$$B(x_0, r) \cap L = \{x \in L \mid ||x - x_0|| < r\} .$$

Sia $x \in B(x_0, r) \cap L$. Essendo x_0, x nel sottospazio L, $y = x - x_0$ è in L e verifica $||y|| < r$; ossia,

$$x = x_0 + y , \qquad y \in B(0, r) \cap L .$$

Dunque, $B(x_0, r) \cap L \subseteq x_0 + B(0, r) \cap L$. L'inclusione opposta si vede in modo analogo. ∎

Siano x_0 ed x_1 due punti di uno s.l.n. X. Il SEGMENTO di estremi x_0 ed x_1 è per definizione l'insieme dei punti

$$x = tx_0 + (1 - t)x_1 , \qquad t \in [0, 1] .$$

(talvolta si considera anche il caso estremo $x_0 = x_1$. In tal caso il segmento si riduce al punto x_0).

Sia $K \subseteq X$ un insieme non vuoto. Si dice che K è CONVESSO quando K contiene un unico punto oppure quando il segmento che unisce due *qualsiasi* punti di K è contenuto in K; ossia quando

$$x_0 , \ x_1 \in K , \quad t \in [0, 1] \Longrightarrow tx_0 + (1 - t)x_1 \in K .$$

Ovviamente, ogni palla è un insieme convesso. Invece, la SUPERFICIE sferica

$$S(x_0, r) = \{x \mid ||x - x_0|| = r\}$$

non è convessa.

Ricordiamo che gli insiemi convessi di $\mathbb{R}$ sono tutti e soli gli intervalli (limitati o meno).

Il risultato seguente è di ovvia dimostrazione:

Teorema 68 *Sia A lineare da X in Y. Se $K \subseteq \mathrm{dom} A$ è convesso in X, allora AK è convesso in Y.*

Le palle centrate in 0 sono insiemi convessi che hanno in più una proprietà di simmetria: se $x \in B(0,r)$ allora $\alpha x \in B(0,r)$ per ogni α tale che $|\alpha| \le 1$.

In generale, un insieme K si dice EQUILIBRATO se

$$|\alpha| \le 1\,, \quad x \in K \quad \Longrightarrow \quad \alpha x \in K\,.$$

Si dice che $\tilde{K}$ è equilibrato rispetto ad un suo punto x_0 se

$$\tilde{K} = x_0 + K\,, \quad \text{con } K \text{ insieme equilibrato.}$$

Si vede immediatamente:

- un insieme di $\mathbb{C}$ equilibrato rispetto a z_0 e che contiene z' contiene anche il disco di centro z_0 e raggio $|z' - z_0|$. Affermazione analoga vale per gli insiemi equilibrati di $\mathbb{R}$, sostituendo i dischi con gli intervalli simmetrici rispetto a z_0. In particolare:

  Lemma 69 *Un insieme equilibrato di $\mathbb{R}$ oppure di $\mathbb{C}$ che è illimitato è uguale, rispettivamente, a $\mathbb{R}$ oppure a $\mathbb{C}$.*

- Se A è un operatore lineare da X in Y e se K è equilibrato in X, allora AK è equilibrato in Y; se K è equilibrato rispetto ad x_0 allora AK è equilibrato rispetto ad Ax_0.

Torniamo a considerare un operatore lineare A da uno s.l.n. X in un s.l.n. Y e consideriamo le palle $B(0,1)$, $B(0,r')$ e $B(x_0,r)$ di X. Supponiamo di sapere che $x_0 \in \mathrm{dom}\, A$. Si ha[11]:

$$AB(0,r) = r(AB(0,1))\,, \qquad AB(x_0,r) = Ax_0 + r(AB(0,1))\,.$$

Di conseguenza,

Lemma 70 *Un operatore lineare A che è limitato su una palla è limitato su ogni altra palla.*

[11] ovviamente, se K è un s.insieme di uno spazio lineare X, di definisce

$$rK = \{rx \,:\, x \in K\}\,.$$

D'altra parte, con $S(0,1) = \{x \mid ||x|| = 1\}$,

$$\text{se } x \in B(0,r) \text{ allora } x = ||x|| \left(\frac{x}{||x||}\right) \text{ con } \frac{x}{||x||} \in S(0,1).$$

Dunque:

Lemma 71 *Vale:*

$$\sup_{||x||_X \le 1} \{||Ax||_Y\} = \sup_{||x||_X = 1} \{||Ax||_Y\} .$$

In particolare, un operatore lineare A è limitato su una palla se e solo se è limitato sulla sfera $S(0,1)$.

Un operatore lineare che è limitato su una palla, non può "crescere troppo velocemente". Infatti:

Teorema 72 *Sia A lineare da X in Y e sia*

$$M_A = \sup_{||x||_X \le 1} \{||Ax||_Y\} < +\infty .$$

Per ogni $x \in X$ vale

$$||Ax||_Y \le M_A ||x||_X . \tag{2.19}$$

Viceversa, se esiste M tale che $||Ax||_Y \le M||x||_X$ allora l'operatore lineare A è limitato su ogni palla.

Dim. Il viceversa è ovvio e quindi basta provare che se A è limitato su $B(0,1)$, allora vale la disuguaglianza (2.19). Basta considerare il caso $x \ne 0$. Se $x \ne 0$, $x/||x|| \in B(0,1)$ e quindi

$$\frac{1}{||x||}||Ax|| = \left|\left|A\frac{x}{||x||}\right|\right| \le M_A .$$

Questa è la disuguaglianza cercata. ∎

Dunque, se A è limitato su una palla, per esso vale

$$\lim_{||x||_X \to 0} ||Ax||_Y = 0$$

ed è quindi continuo in 0. Viceversa, il teorema della limitatezza locale (applicato alla funzione a valori reali $x \to ||Ax||$) mostra che se A è continuo in 0 allora è limitato su una palla $B(0,r)$, e quindi su ogni altra palla. Vale dunque:

Teorema 73 *Un operatore lineare A da X in Y è continuo in* 0 *se e solo se è limitato su una* qualsiasi *palla; equivalentemente, se e solo se esiste un numero M per cui vale*

$$||Ax||_Y \leq M||x||_X \, .$$

Ciò suggerisce un punto di partenza per lo studio della continuità degli operatori lineari. Prima di fare ciò, ricordiamo che in generale si chiama "funzione limitata" una funzione che ha immagine limitata. Il Teorema 66 mostra che l'unico operatore lineare "limitato", con immagine limitata, da X in Y è quello identicamente zero; e quindi il termine "limitato" riferito ad operatori lineari rimane libero, e può essere usato con un significato diverso. Chiamiamo quindi operatore LINEARE LIMITATO un operatore lineare che è limitato su una (qualsiasi) palla; ossia un operatore lineare per il quale vale la disuguaglianza (2.19). Dunque:

Teorema 74 *Un operatore lineare A da X in Y è continuo in* 0 *se e solo se è limitato.*

2.6.2 La continuità degli operatori lineari

E' noto:

Teorema 75 *Se* $\dim X < +\infty$ *e se la trasformazione A è lineare da X in Y allora A è continua.*

Gli esempi seguenti mostrano che, se X ha dimensione infinita, allora esistono sia operatori lineari continui che non continui. Un esempio banale di operatore lineare continuo è quello che ad ogni elemento di X associa lo 0 di Y, ossia il funzionale nullo. Un esempio meno banale è il seguente:

Esempio 76 Sia $X = C(a, b)$, $Y = \Phi$ e sia

$$\operatorname{dom} A = X \, , \qquad Ax = x(a) \, .$$

Essendo

$$||Ax - Ay||_\Phi = |x(a) - y(a)| \leq ||x - y||_X$$

si vede che A è addirittura *uniformemente continuo.* ∎

Mostriamo ora un esempio di *funzionale lineare non continuo.*

Esempio 77 Su $C(-1,1)$ si consideri l'operatore lineare Ψ definito da

$$\mathrm{dom}\Psi = \{x \text{ derivabili in } 1\}\,, \qquad \Psi x = x'(1)\,.$$

Quest'operatore, chiaramente lineare, *non è continuo.* Per mostrare ciò si consideri la successione delle funzioni x_n

$$x_n(t) = t^n/n\,.$$

Da $||x_n|| < 1/n$ segue che

$$\lim x_n = x_0 = 0$$

mentre per ogni n si ha:

$$\Psi x_n = 1\,;$$

ossia,

$$\lim \Psi x_n \neq \Psi x_0$$

e quindi Ψ non è continuo. ∎

Si noti che l'esempio precedente mostra che anche *funzionali lineari importanti per le applicazioni possono essere discontinui;* e, l'esempio specifico spiega perchè il problema della *derivazione numerica* è assai delicato.

Esempi di operatori lineari, rispettivamente continui e non continui, tra s.l.n-ti ambedue di dimensione infinita sono i seguenti:

Esempio 78 Sia $X = Y = C(a,b)$ e sia A con dominio uguale ad X,

$$(Ax)(t) = \int_a^t x(s)\,\mathrm{d}s\,.$$

E'

$$||Ax - Ay||_Y = \max_{[a,b]} \left| \int_a^t [x(s) - y(s)]\,\mathrm{d}s \right| \leq (b-a) \cdot ||x - y||_X\,;$$

e nuovamente si vede che l'operatore A è *uniformemente continuo.*

Un esempio di operatore *lineare discontinuo* è il seguente:

Sia $X = L^2(0,1)$, $Y = C([0,1])$. Il dominio di A sia lo spazio lineare delle classi di equivalenza dotate di rappresentante continuo. Se x è una funzione continua, indichiamo con $[x]$ la sua classe di equivalenza. Si noti che il simbolo non è ambiguo perché una classe di equivalenza di $L^2(0,1)$ contiene *al più* una funzione continua. Sia

$$A[x] = y\,, \quad \text{ove} \quad y(t) \equiv x(1)\,.$$

Le (classi di equivalenza delle) funzioni

$$x_n(t) = t^n$$

costituiscono una successione in $L^2(0,1)$, convergente a 0; mentre, per ogni x,

$$A[x_n] \equiv 1 \,. \quad \blacksquare$$

Passiamo ora a studiare le proprietà degli operatori lineari che sono anche continui. Abbiamo notato che gli operatori lineari continui degli esempi precedenti sono anche uniformemente continui. Come ora vedremo, è questo un fatto generale.

Teorema 79 *Siano X ed Y due s.l.n-ti e sia A lineare da X in Y. Vale:*

- *l'operatore A è continuo in ciascun punto del suo dominio se e solo se è continuo in un punto;*
- *l'operatore A è continuo se e solo se è uniformemente continuo;*
- *l'operatore A è continuo se e solo se è limitato.*

Dim. Proviamo che se A è continuo in un punto x_0 allora esso è continuo in qualunque altro punto x_1. Ovviamente, sia x_0 che x_1 devono appartenere al dominio di A. Sia $\epsilon > 0$ e sia $B(x_0, \delta)$ tale che

$$x \in B(x_0, \delta) \cap \mathrm{dom} A \Longrightarrow ||Ax - Ax_0|| < \epsilon \,. \tag{2.20}$$

Sia ora $x \in B(x_1, \delta) \cap \mathrm{dom} A$. Usando il Lemma 67 si vede che x può rappresentarsi come

$$x = x_1 + (x' - x_0) \,, \qquad \text{con} \quad x' = (x - x_1) + x_0 \in \mathrm{dom} A \cap B(x_0, \delta) \,.$$

Dunque,

$$||Ax - Ax_1|| = ||A(x_1 - (x' - x_0)) - Ax_1|| = ||A(x' - x_0)|| = ||Ax' - Ax_0|| < \epsilon \,.$$

L'ultima disuguaglianza segue da (2.20).

Ciò prova la continuità in x_1 e prova anche che il numero δ nel punto x_1 è il medesimo usato in x_0. Essendo x_1 arbitrario, si ha la *continuità uniforme.*

In particolare, la continuità in un qualsiasi punto x_0 equivale alla continuità in 0, e quindi alla limitatezza, si veda il Teorema 74. ∎

Frequentemente conviene verificare la continuità di un operatore verificando direttamente che è limitato.

Come si è visto, l'operatore lineare A è continuo se e solo se

$$M_A = \sup_{||x||_X \leq 1} ||Ax||_Y < +\infty \,.$$

Ci si può chiedere se l'estremo superiore sia in realtà un massimo. E' facile immaginare che l'estremo superiore non sarà un massimo se il dominio di A non è chiuso. Però:

Teorema 80 *Esistono s.l.n-ti* completi X *ed operatori lineari* continui A *con dominio* uguale *ad* X *e tali che*

$$\max\{||Ax||_Y \mid ||x||_X \leq 1\}$$

non *esiste.*

Dim. Si scelga $X = L^1(0,1)$ ed il funzionale

$$Lx = \int_0^1 sx(s)\,\mathrm{d}s\,.$$

E' immediato verificare che questo funzionale è continuo e che

$$||x|| \leq 1 \quad \Longrightarrow \quad |Lx| \leq 1\,;$$

anzi, si vede che

$$\sup\{||Lx||_Y\,,\ \mid ||x||_X \leq 1\} = 1\,.$$

Infatti, se

$$x_n(t) = \begin{cases} 0 & \text{se} \quad 0 \leq t \leq 1 - 1/n \\ n & \text{se} \quad 1 - \frac{1}{n} \leq t \leq 1 \end{cases}$$

allora:

$$\lim Lx_n = 1\,.$$

Mostriamo che se $||x||_{L^1(0,1)} \leq 1$ allora *non* vale $Lx = 1$. Sia infatti x tale che $Lx = 1$. In questo caso x non è zero q.o. e quindi esiste $\delta \in (0,1)$ tale che

$$\int_0^\delta |x(s)|\,\mathrm{d}s = \alpha > 0\,.$$

sottolineiamo che il numero δ *si può scegliere* minore *di* 1. Si scriva ora:

$$\begin{aligned} 1 &= \int_0^1 sx(s)\,\mathrm{d}s = \int_0^\delta sx(s)\,\mathrm{d}s + \int_\delta^1 sx(s)\,\mathrm{d}s \\ &\leq \delta \int_0^\delta |x(s)|\,\mathrm{d}s + \int_\delta^1 |x(s)|\,\mathrm{d}s \\ &= \int_0^1 |x(s)|\,\mathrm{d}s - (1-\delta) \int_0^\delta |x(s)|\,\mathrm{d}s \end{aligned}$$

ossia

$$\int_0^1 |x(s)|\,\mathrm{d}s \geq 1 + (1-\delta)\alpha\,.$$

Dunque, per tale funzione x si ha $||x|| > 1$ e quindi il massimo sull'insieme $\{x \mid ||x|| \leq 1\}$ non viene raggiunto. ∎

Infine, introduciamo due operatori particolari, ed i loro simboli: col simbolo 0, riferito ad operatori che operano da X in Y, si intende l'OPERATORE NULLO, ossia quello che associa ad ogni $x \in X$ l'elemento 0 di Y. Col simbolo I, riferito ad operatori da X in X, si intende l'OPERATORE IDENTITÀ, ossia quell'operatore che ad ogni x di X associa se stesso:

$$Ix = x\,.$$

Osservazione 81 I punti che non appartengono al dominio di un operatore non influiscono sulla proprietà di continuità. Quindi, sostituendo X con $\mathrm{dom}\,A$, avremmo potuto assumere che l'operatore A fosse definito su tutto lo spazio, evitando alcune complicazioni. Abbiamo preferito la via precedente perché di fatto molti degli operatori che si incontrano in pratica hanno solamente dominio denso e talvolta è effettivamente necessario tenere conto di ciò, si veda il Teorema 87. ∎

2.6.3 Funzionali lineari continui ed iperpiani

Proviamo:

Teorema 82 *Sia* Ψ *un* funzionale *lineare definito su* X *non identicamente nullo, continuo o meno. Il suo nucleo ammette complementare di dimensione* 1.

Dim. Sia x_0 tale che $\Psi x_0 \neq 0$. Si noti che per ogni $x \in X$

$$n_x = x - x_0 \frac{\Psi x}{\Psi x_0} \in \ker \Psi\,.$$

Ogni $x \in X$ si rappresenta come

$$x = n_x + \alpha_x x_0\,, \qquad \alpha_x = \frac{\Psi x}{\Psi x_0}\,.$$

Questa rappresentazione è unica perché se si ha anche

$$x = n'_x + \alpha'_x x_0$$

allora sottraendo si trova

$$0 = (n_x - n'_x) + (\alpha_x - \alpha'_x)x_0 \,.$$

Applicando il funzionale Ψ ai due membri si trova

$$0 = (\alpha_x - \alpha'_x)\Psi x_0$$

e quindi $\alpha_x = \alpha'_x$, perché $\Psi x_0 \neq 0$. E dunque si ha anche $n_x = n'_x$.

Ciò prova che $\{\beta x_0 \,,\ \beta \in \Phi\}$ è uno spazio complementare di $\ker \Psi$. ∎

Osservazione 83 Il teorema precedente non richiede la completezza di X e nemmeno richiede la chiusura di $\ker \Psi$. Ciò nonostante, asserisce che $\ker \Psi$ ha un complementare di dimensione 1, e quindi *chiuso.* Mostriamo su un esempio la costruzione di tale complementare, nel caso in cui $\ker \Psi$ non è chiuso. Sia X lo s.l.n. delle funzioni continue su $[-1, 1]$ e derivabili su $(-1, 1)$, dotato della norma del massimo. Questo spazio non è completo. Sia

$$\Psi x = x'(0) \,.$$

Il funzionale Ψ è ovunque definito, e non è continuo, come si vede facilmente riadattando gli argomenti presentati nell'esempio 77.

Il nucleo di Ψ è l'insieme delle funzioni di $C^1(-1, 1)$ la cui derivata è nulla in 0. Non è difficile mostrare che questo spazio lineare è denso in $C(-1, 1)$ e quindi in X. e quindi $\ker \Psi$ non è chiuso in X. Ciò nonostante ammette complementare chiuso: ogni $x \in X$ si rappresenta in modo unico come

$$x(t) = [x(t) - x'(0)t] + x'(0)t$$

somma di un elemento di $\ker \Psi$ e di un multiplo di x_0, $x_0(t) = t \notin \ker \Psi$. ∎

Se Ψ è una qualsiasi trasformazione continua tra s.l.n-ti X ed Y, l'insieme degli zeri di Ψ è chiuso, come controimmagine continua di un chiuso. *Il viceversa vale nel caso particolare dei funzionali lineari:*

Teorema 84 *Sia Ψ un* funzionale *lineare su uno s.l.n. X. Esso è continuo se e solo se il suo nucleo è chiuso.*

Dim. Se il nucleo di Ψ è tutto X allora Ψ è costante (uguale a 0) e quindi è continuo. Altrimenti, sia $x_0 \notin \ker \Psi$. Dato che $\ker \Psi$ è chiuso e diverso da X, esiste $\delta > 0$ tale che

$$\delta = \operatorname{dist}(x_0, \ker \Psi) = \inf\{||x - x_0|| \,,\ x \in \ker \Psi\} \,.$$

Sia $B(x_0, \delta/2) = \{x \mid ||x - x_0|| < \delta/2\}$. L'immagine $\Psi B(x_0, \delta/2)$ di $B(x_0, \delta/2)$ è un insieme equilibrato (rispetto a $\Psi(x_0)$) in $\mathbb{R}$ oppure in $\mathbb{C}$, che non contiene 0, perché $B(x_0, \delta/2)$ non interseca $\ker \Psi$. Per il Lemma 69, $\Psi B(x_0, \delta/2)$ è limitato, e quindi Ψ è continuo. ∎

Si chiamano IPERPIANI i sottospazi chiusi di codimensione 1 e gli insiemi che si ottengono da essi per traslazione. Dunque:

Teorema 85 *Gli iperpiani sono tutti e soli gli insiemi della forma*

$$H = \{x \mid \Psi x = c\}$$

ove Ψ è un funzionale lineare e continuo.

Dim. Se ψ è un funzionale lineare continuo, il suo nucleo è chiuso, come contrimmagine dell'insieme chiuso $\{0\}$.

Viceversa, sia N un spazio lineare chiuso di codimensione 1. Costruiamo un funzionale lineare Ψ che ha N per nucleo, e che quindi è continuo per il Teorema 84.

Essendo N di codimensione 1, esiste $x_0 \notin N$ tale che ogni elemento di X si rappresenta in modo unico come

$$x = n_x + \alpha_x x_0 \,, \qquad n \in N \,.$$

Il funzionale cercato è quello che ad x associa il numero α_x. ∎

Ossia, *gli iperpiani sono gli insiemi di livello di funzionali lineari e continui.*

Se Ψ è un funzionale lineare continuo, definiamo i due SEMISPAZI

$$H_+ = \{x \mid \Psi(x) > c\} \,, \qquad H_- = \{x \mid \Psi(x) < c\} \,.$$

I due semispazi H_+ ed H_- sono ovviamente disgiunti (perché le disuguaglianze sono strette). Le loro chiusure, che si chiamano anche semispazi chiusi, hanno in comune i punti dell'iperpiano $\{x \mid \Psi(x) = c\}$.

E' opportuno notare che le notazioni H_+ ed H_- non hanno significato intrinseco. Infatti, il funzionale Ψ che il teorema 85 associa ad H non è unico. Se, per esempio, $c = 0$, allora si identifica lo stesso iperpiano H sia col funzionale Ψ che col funzionale $-\Psi$; e lo scambio di Ψ con $-\Psi$ scambia tra di loro i due semispazi.

Notiamo infine che il teorema 84 vale per i *funzionali.* Non vale per generici operatori lineari, come mostra l'esempio seguente.

Esempio 86 Sia $X = Y = C(0,1)$ e sia

$$\operatorname{dom} A = C^1(0,1)\,, \qquad Ax = x'\,.$$

Argomenti analoghi a quelli visti all'esempio 77 mostrano che A non è continuo. Il suo nucleo è il s.spazio i cui elementi sono le funzioni costanti, e quindi è chiuso nonostante che l'operatore A *non è continuo.* ∎

2.6.3.1 Notazioni

In pratica quando Ψ è un funzionale lineare e continuo su X, invece di scrivere $\Psi(x)$ per indicare il valore preso da Ψ nel punto x, si scrive

$$\langle\!\langle \Psi, x\rangle\!\rangle$$

ossia

$$\langle\!\langle \Psi, x\rangle\!\rangle = \Psi(x)\,.$$

Avremo modo di vedere la comodità di questa notazione. Va notato esplicitamente che il simbolo del funzionale lineare viene scritto sulla *sinistra,* così come sulla sinistra compare nella notazione $\Psi(x)$. Insistiamo su questo perché in alcuni libri esso si trova scritto sulla *destra,* ossia si trova scritto $\langle\!\langle x, \Psi\rangle\!\rangle$ invece di $\langle\!\langle \Psi, x\rangle\!\rangle$. Ci sono buone ragioni per l'uso di ambedue le notazioni[12].

Lo spazio lineare di tutti i *funzionali lineari e continui* definiti su X si chiama lo SPAZIO DUALE di X, e si indica col simbolo X^* oppure X'.

2.6.4 Lo spazio $\mathcal{L}(X,Y)$

Siano X ed Y due s.l.n-ti ed A, B due operatori lineari da X in Y. Definendo

$$\operatorname{dom}(A+B) = (\operatorname{dom}A)\cap(\operatorname{dom}B)\,, \qquad (A+B)x = Ax + Bx\,,$$

si ottiene chiaramente un operatore lineare $A+B$; ma in generale $\operatorname{dom}(A+B)$, $\operatorname{dom}A$ e $\operatorname{dom}B$ sono diversi e quindi non è possibile dare una struttura di spazio lineare all'insieme di tutti gli operatori lineari da X in Y. Per esempio, $B + (-B)$ non è in generale l'operatore 0, perchè l'operatore 0 è definito su X mentre $B + (-B)$ è solo definito su $\operatorname{dom}B$; e quindi $A + B + (-B)$ non è, in generale, l'operatore A. Se però ci si limita a considerare soltanto gli operatori lineari e continui si può ottenere di più. Vale infatti:

[12] In realtà la notazione comunemente usata è $\langle\cdot,\cdot\rangle$. Noi usiamo la notazione $\langle\!\langle\cdot,\cdot\rangle\!\rangle$ perché la notazione $\langle\cdot,\cdot\rangle$ si usa anche per indicare i "prodotti interni" nel contesto degli spazi di Hilbert. Dato che vedremo una relazione tra funzionali lineari e prodotti interni, è opportuno essere precisi nel distinguere gli uni dagli altri.

Teorema 87 *Sia A un operatore lineare e continuo da X in Y. Se Y è* completo *allora l'operatore A ammette un'unica estensione continua alla chiusura del suo dominio.*

Dim. Presentiamo i punti salienti della dimostrazione (del tutto analoga a quella che si usa per costruire l'estensione per continuità di funzioni reali), per mostrare il ruolo della completezza di Y.

Se x_0 è un punto della chiusura del dominio di A, esiste una successione (x_n) convergente ad x_0, $x_n \in \mathrm{dom} A$ (si noti che se $x_0 \in \mathrm{dom} A$ allora si può scegliere $x_n = x_0$ per ogni n).

Per la continuità di A si ottiene

$$||Ax_n - Ax_m||_Y = ||A(x_n - x_m)||_Y \leq M||x_n - x_m||_X \,. \tag{2.21}$$

La successione (x_n) è fondamentale in X, essendo per ipotesi convergente. Dunque, anche la successione (y_n), $y_n = Ax_n$ è fondamentale, però nello spazio Y. Essendo Y completo, si ha

$$\lim Ax_n = y_0 \,.$$

Si definisce quindi

$$\tilde{A}x_0 = y_0.$$

Se (x'_n) è una seconda successione in $rmdom\, A$ convergente ad x_0, vale

$$||Ax_n - Ax'_n||_Y \leq M||x_n - x'_n||_X \to 0$$

e quindi

$$\lim Ax_n = \lim Ax'_n \,;$$

ossia il valore y_0 dipende solo da x_0, e non dalla particolare successione scelta per calcolarlo. Dunque l'operatore $\tilde{A}$ che abbiamo costruito è un operatore univoco.

Ovviamente, $\tilde{A}$ estende A: se $x_0 \in \mathrm{dom} A$, scegliendo $x_n = x_0$ per ogni n si vede che

$$\tilde{A}x_0 = Ax_0 \,.$$

Si prova facilmente che l'operatore $\tilde{A}$ è lineare, ed è limitato.

Lasciamo per esercizio la dimostrazione della linearità e proviamo la limitatezza: se $x_n \to x_0$ ed $Ax_n \to y_0$,

$$||\tilde{A}x_0||_Y = \lim ||Ax_n||_Y \leq M \lim ||x_n||_X = M \cdot ||x_0||_X \tag{2.22}$$

(si ricordi che la norma è continua). ∎

Ricapitolando, $\tilde{A}$ è (l'unica) estensione continua di A alla chiusura del suo dominio. In particolare, se il dominio di A è denso in X, allora $\tilde{A}$ è definito su X.

Naturalmente, in pratica identificheremo A con $\tilde{A}$ (usando il simbolo più semplice A per ambedue gli operatori).

Da ora in poi, lavorando con operatori lineari e continui a valori in uno spazio completo, assumeremo di averli estesi per continuità alla chiusura del dominio; e se non diversamente detto, assumeremo che il dominio sia X. Lavorando con operatori definiti su X, sia $A+B$ che αA (definito da $(\alpha A)x = \alpha(Ax)$ per ogni x) hanno dominio X e sono continui. Dunque, l'insieme degli operatori lineari continui su X è uno spazio lineare. Ciò che è più importante, *esso può essere dotato di norma,* come segue:

$$||A|| = \sup_{||x||_X \leq 1} ||Ax||_Y \,. \tag{2.23}$$

Si lascia al lettore la facile verifica che quella appena definita è una norma.

Conviene notare una conseguenza utile della definizione (2.23):

Corollario 88 *Per ogni $x \in X$ vale:*

$$||Ax||_Y \leq ||A|| \cdot ||x||_X \,. \tag{2.24}$$

Se Z è uno s.l.n. completo, e B è lineare e continuo da Y a Z, allora vale

$$||BA|| \leq ||B|| \cdot ||A|| \,. \tag{2.25}$$

Dim. Infatti, se $||\xi||_X \leq 1$, allora vale $||A\xi||_Y \leq ||A||$. Se $x \neq 0$ allora $\xi = x/||x||_X$ ha norma 1 e quindi

$$||A|| \geq ||A\xi||_Y = \frac{1}{||x||_X} \cdot ||Ax||_Y$$

ossia la (2.24).

La (2.24) mostra che:

$$||BAx|| \leq ||B|| \cdot ||Ax||_Y \leq ||B|| \cdot ||A|| \cdot ||x||_X \,.$$

Prendendo l'estremo superiore per $||x||_X \leq 1$, si trova la (2.25). ∎

La disuguaglianza (2.25) nel caso in cui $Z = Y = X$ mostra

$$||A^2|| \leq ||A||^2$$

e, più in generale,

$$||A^n|| \leq ||A||^n \,.$$

Conviene mostrare subito un modo equivalente per il calcolo di $||A||$:

Teorema 89 *Vale:*

$$||A|| = \min\{M \mid ||Ax||_Y \le M||x||\}\,. \tag{2.26}$$

<u>**Dim.**</u> Indichiamo con M_0 il numero

$$M_0 = \inf\{M \mid ||Ax||_Y \le M||x||\}$$

e proviamo che $M_0 = ||A||$, ossia che

$$M_0 = \sup_{||x||_X \le 1} ||Ax||_Y\,.$$

Ciò in particolare mostra che l'estremo inferiore è un minimo.

La (2.24) implica che $M_0 \le ||A||$. Per mostrare la disuguaglianza opposta, fissiamo $\delta > 0$ arbitrario. Vale, per ogni x,

$$||Ax||_Y \le (M_0 + \delta) \cdot ||x||_X$$

e quindi, se $||x||_X \le 1$,

$$||Ax||_Y \le (M_0 + \delta) \cdot ||x||_X \le M_0 + \delta\,.$$

Dunque, la disuguaglianza

$$||A|| \le M_0 + \delta$$

vale per ogni $\delta > 0$. Passando all'estremo inferiore rispetto a δ si trova

$$||A|| \le M_0 \tag{2.27}$$

e quindi l'uguaglianza (2.26). ∎

Possiamo ora tornare a considerare la diseguaglianza (2.22). Essa può ora scriversi

$$||\tilde{A}x_0|| \le M||x_0|| \qquad M = \sup_{||x|| \le 1\ x \in \mathrm{dom} A} ||Ax||$$

ossia

$$||\tilde{A}|| \le M\,.$$

Però, $\tilde{A}$ estende A e quindi $||\tilde{A}|| \ge M$. dunque:

Corollario 90 *La norma dell'operatore $\tilde{A}$, estensione per continuità di A (si veda il Teorema 87), è uguale alla norma di A:*

$$||\tilde{A}|| = \sup_{||x|| \le 1} ||\tilde{A}x|| = \sup_{||x|| \le 1\ x \in \mathrm{dom} A} ||Ax|| = ||A||\,.$$

Ossia, *il calcolo della norma di un operatore lineare continuo definito su X può effettuarsi a partire da una sua restrizione ad un sottospazio denso in X.*

Si lascia per esercizio di provare la seguente ulteriore caratterizzazione di $||A||$:

Teorema 91 *Vale:*

$$||A|| = \sup_{||x||=1} ||Ax||_Y = \sup_{x\neq 0} \frac{||Ax||_Y}{||x||_X} .$$

Quando sia X che Y sono s.l.n-ti completi, ossia *spazi di Banach,* lo spazio degli operatori lineari e continui da X in Y, normato nel modo che abbiamo appena introdotto, si indica col simbolo $\boxed{\mathcal{L}(X,Y)}$ oppure $\boxed{\mathcal{B}(X,Y)}$. Due casi sono di uso particolarmente frequente e ad essi si riservano simboli speciali: il caso in cui $X = Y$ ed il caso, importantissimo, $X = \Phi$. Nel primo caso si usa il simbolo $\boxed{\mathcal{L}(X)}$ invece di $\mathcal{L}(X,X)$; nel secondo caso, come si è già detto, si usa il simbolo $\boxed{X^*}$ o $\boxed{X'}$ invece di $\mathcal{L}(X,\Phi)$. Lo spazio X^* si chiama lo SPAZIO DUALE di X.

Infine, esaminiamo il problema della completezza dello spazio $\mathcal{L}(X,Y)$. La dimostrazione del teorema seguente usa la completezza dello spazio Y ma non quella dello spazio X. Per questa dimostrazione abbiamo bisogno di ricordare che una successione fondamentale è anche limitata; e che la successione (A_n) è limitata in $\mathcal{L}(X,Y)$ quando esiste un numero M, indipendente da n, tale che

$$||A_n|| < M .$$

Teorema 92 *Lo spazio $\mathcal{L}(X,Y)$ è completo.*

<u>**Dim**</u>**.** Dobbiamo mostrare che ogni successione (A_n) fondamentale in $\mathcal{L}(X,Y)$ è anche convergente. Sia allora (A_n) fondamentale. Per definizione di norma in $\mathcal{L}(X,Y)$, per ogni $\epsilon > 0$ esiste N_ϵ tale che per n, m maggiori di N_ϵ vale

$$||A_n - A_m|| \le \epsilon \qquad \text{ossia} \qquad \sup_{||x||_X \le 1} ||(A_n - A_m)x||_Y \le \epsilon . \tag{2.28}$$

Segue che la successione $(A_n x)$ di elementi di Y è fondamentale per ogni x di norma minore o uguale ad 1; e quindi per ogni $x \in X$. Ciò permette di definire l'operatore B dato da

$$Bx = \lim A_n x .$$

Proviamo la linearità e la continuità di B e poi proviamo che $\lim A_n = B$. Da

$$B(\alpha x + \beta y) = \lim A_n(\alpha x + \beta y) = \lim \{\alpha A_n x + \beta A_n y\} = \alpha Bx + \beta By$$

segue la linearità. La continuità segue perché, se $||x|| \leq 1$,

$$||Bx|| = ||\lim A_n x|| = \lim ||A_n x|| \leq M||x||$$

con M *indipendente da* n *perchè la successione* (A_n), *essendo fondamentale, è limitata*[13].

Mostriamo ora che $B = \lim A_n$, ossia che

$$\lim ||B - A_n|| = 0\,. \tag{2.29}$$

E':

$$||B - A_n|| = \sup_{||x|| \leq 1} ||(B - A_n)x||_Y\,.$$

Sia $\epsilon > 0$ e sia N_ϵ tale che, per n, m maggiori di N_ϵ, valga (2.28). Fissato x con $||x|| < 1$, scriviamo

$$||Bx - A_n x||_Y = ||(B - A_m)x + (A_m - A_n)x||_Y \leq ||(B - A_m)x||_Y + \epsilon\,. \tag{2.30}$$

Notiamo che questa disuguaglianza vale per ogni x *con* $||x|| \leq 1$ *e per tutti gli* $m > N_\epsilon$.

La definizione di B mostra l'esistenza di un opportuno $m > N_\epsilon$ (dipendente sia da x che da ϵ) per cui vale anche

$$||(B - A_m)x|| < \epsilon\,.$$

Dunque, la (2.30) dà:

$$||Bx - A_n x||_Y \leq \inf_m \{||(B - A_m)x||_Y + \epsilon\} \leq 2\epsilon \qquad \forall n > N_\epsilon\,.$$

Ciò prova (2.29). ∎

2.7 Esempi di spazi duali

Sia X uno spazio di Banach. Per definizione, X^* è lo spazio (di Banach per il Teorema 92) dei funzionali lineari e continui su X. Il problema che vogliamo studiare ora è il seguente: se X è uno spazio "particolare", per esempio uno spazio di funzioni o di successioni, vogliamo vedere se esiste un altro spazio "particolare" Y che è *isometricamente isomorfo* ad X^*; ossia tale che esista una trasformazione L da Y in X^* che è 1) suriettiva; 2) isometrica;

[13] si noti che in questa dimostrazione si usa anche la continuità della norma.

3) lineare (se $\Phi = \mathbb{R}$) oppure antilineare (se $\Phi = \mathbb{C}$).[14] Ricordiamo che L è una TRASFORMAZIONE ISOMETRICA quando "conserva la norma", ossia quando

$$||Ly||_{X^*} = ||y||_Y \, .$$

E quindi una trasformazione isometrica è necessariamente iniettiva.

Se L esiste, Y viene ad avere tutte le proprietà metriche di X^*. Si dice allora che Y è una REALIZZAZIONE di X^* e, frequentemente, non si distingue tra Y ed X^*.

Un esempio particolare è ben noto: se $X = l^2(n)$, lo spazio euclideo n-dimensionale, allora una realizzazione del duale è lo spazio stesso.

Non sempre è possibile trovare delle realizzazioni concrete (e comode) di uno spazio duale; e d'altra parte esistono spazi di Banach che non sono isometricamente isomorfi a nessuno spazio duale. Per questo conviene elencare alcuni casi particolarmente importanti. Prima di presentare le dimostrazioni, raccogliamo i risultati nella tabella seguente:

spazio	duale
l^1	l^∞
$l^p \, , \quad p < +\infty$	$l^{p'} \, , \quad p' = p/(p-1)$
c_0	l^1
$L^1(\Omega)$	$L^\infty(\Omega)$
$L^p(\Omega) \, , \quad p < +\infty$	$L^{p'}(\Omega) \, , \quad p' = p/(p-1)$
$C(a,b)$	$NV(a,b)$

Lo spazio $NV(a,b)$ è definito in seguito.

Non abbiamo inserito nella tabella precedente gli spazi l^∞ ed $L^\infty(\Omega)$. Realizzazioni dei loro duali sono note, ma per descriverle avremmo bisogno di conoscenze di teoria della misura che non abbiamo presentato.

2.7.0.1 Il duale di l^p, $1 \le p < +\infty$

Per caratterizzare il duale di l^p ed anche di c_0 abbiamo bisogno di una particolare successione di elementi dello spazio l^p stesso, che indichiamo con $(e^{(n)})$. Dunque, ciascun $e^{(n)}$ è a sua volta una successione di numeri. Per definizione,

$$e_i^{(n)} = \begin{cases} 1 & \text{se } i = n \\ 0 & \text{altrimenti.} \end{cases} \qquad (2.31)$$

[14]Una trasformazione L si dice ANTILINEARE se vale $L(\alpha x + \beta y) = \bar{\alpha} Lx + \bar{\beta} Ly$.

Notiamo che $||e^{(n)}||_p = 1$ per ogni p, $1 \le p \le +\infty$ e che lo spazio lineare generato dagli elementi $e^{(n)}$ è denso in l^p per ogni p, $1 \le p < +\infty$. *Non è invece denso in* l^∞*.*

Sia $X = l^p$ con $1 \le p < +\infty$. In questo caso una realizzazione di X^* è

$$l^{p'} \qquad \text{con} \qquad p' = \begin{cases} +\infty & \text{se} \quad p = 1 \\ \frac{p}{p-1} & \text{se} \quad p > 1\,. \end{cases} \tag{2.32}$$

Proviamo ciò prima di tutto nel caso $p = 1$. Sia $(y_n) \in L^\infty$. Si vede immediatamente che il funzionale lineare x^*, dipendente da (y_n),

$$\langle\langle x^*, x\rangle\rangle = \sum_{n=0}^{+\infty} \bar{y}_n x_n \tag{2.33}$$

è lineare, ed è continuo perché

$$|\langle\langle x^*, x\rangle\rangle| = |\sum_{n=0}^{+\infty} \bar{y}_n x_n| \le \sup_n\{|y_n|\} \cdot ||x||_1\,.$$

Inoltre, la trasformazione $y = (y_n) \to x^*$ è antilineare e si vede facilmente che è isometrica (e quindi anche iniettiva). Infatti, la disuguaglianza precedente mostra che

$$||x^*|| = \sup_{||x||_1=1} \langle\langle x^*, x\rangle\rangle \le ||y||_\infty\,. \tag{2.34}$$

Per vedere che vale anche la disuguaglianza inversa, e quindi l'uguaglianza, si scelga la successione $x^{(N)}$ definita da

$$x_r^{(N)} = \begin{cases} \frac{y_r}{|y_r|} & \text{se } r = N \text{ e } y_r \ne 0 \\ 0 & \text{altrimenti.} \end{cases}$$

Ovviamente, $||x^{(N)}||_1 \le 1$ e

$$\sup_N \langle\langle x^*, x^{(N)}\rangle\rangle = \sup_N |y_N| = ||y||_\infty$$

e quindi in (2.34) vale l'uguaglianza.

Per completare la dimostrazione, dobbiamo far vedere che la trasformazione che ad $y \in l^\infty$ associa $x^* \in (l^1)^*$ data da (2.33) è suriettiva; ossia dobbiamo assegnare un qualsiasi $x^* \in (l^1)^*$ ed associargli un opportuno $y \in l^\infty$, in modo che valga (2.33). Per costruire y consideriamo la successione $e^{(n)}$ in (2.31). Definiamo y ponendo

$$y_n = \overline{\langle\langle x^*, e^{(n)}\rangle\rangle}\,.$$

Da

$$|y_n| \leq ||x^*||$$

segue che $y \in l^\infty$.

Sia ora $x \in l^1$. Associamogli la successione $x^{(N)}$ definita come segue:

$$x^{(N)} = \sum_{k=0}^{N} x_k e^{(k)} \qquad \text{ossia} \quad x_k^{(N)} = \begin{cases} x_k & \text{se } k \leq N \\ 0 & \text{altrimenti.} \end{cases}$$

Si ha

$$\begin{aligned} \langle\!\langle x^*, x \rangle\!\rangle &= \lim_N \langle\!\langle x^*, x^{(N)} \rangle\!\rangle = \lim_N \langle\!\langle x^*, \sum_{k=0}^{N} x_k e^{(k)} \rangle\!\rangle \\ &= \lim_N \sum_{k=0}^{N} x_k \langle\!\langle x^*, e^{(k)} \rangle\!\rangle = \lim_N \sum_{k=0}^{N} \bar{y}_k x_k = \sum_{k=0}^{+\infty} \bar{y}_k x_k \,. \end{aligned}$$

Si noti che l'ultima uguaglianza si giustifica perchè già sappiamo che $y \in l^\infty$ e già sappiamo che, in tal caso,

$$x \to \sum_{k=0}^{+\infty} \bar{y}_k x_k$$

è continua su l^1.

Ciò completa l'analisi del caso $p = 1$.

In modo analogo **trattiamo il caso** $1 < p < +\infty$.

Siano $x = (x_n) \in l^p$ ed $y = (y_n) \in l^{p'}$. Dalla disuguaglianza di Hölder si vede che:

$$\sum_{k=0}^{+\infty} \bar{y}_n x_n \leq ||(y_n)||_{p'} \cdot ||(x_n)||_p = ||y||_{p'} \cdot ||x||_p \,.$$

Ciò mostra che la trasformazione lineare

$$x \to \sum_{k=0}^{+\infty} \bar{y}_k x_k$$

è continua su l^p e suggerisce di considerare la trasformazione L da $Y = l^{p'}$ in X^*:

$$(Ly)(x) = \sum_{k=0}^{+\infty} \bar{y}_n x_n \,,$$

che è chiaramente iniettiva e inoltre

$$||Ly||_{X^*} \leq ||y||_{p'}\,.$$

Si vede che vale l'uguaglianza. Infatti, sia

$$x_n = \left(|y_n|^{p'/p} \frac{y_n}{|y_n|} \right) \frac{1}{||y||_{p'}^{(p'/p)}} \quad \text{se } y_n \neq 0,\ x_n = 0 \text{ altrimenti.}$$

Si vede facilmente che $x = (x_n)$ è un elemento di $X = l^p$ di norma 1. Per esso vale

$$(Ly)(x) = \frac{1}{||y||_{p'}^{p'/p}} ||y||_{p'}^{p'} = ||y||_{p'}\,.$$

In questo modo si è trovata una trasformazione antilineare L che è isometrica (e quindi anche iniettiva) da $l^{p'}$ in $(l^p)^*$.

Per concludere, basta mostrare che L è suriettiva, ossia che ogni elemento di $(l^p)^*$ si rappresenta come in (2.32). Sia allora $x^* \in (l^p)^*$. Dobbiamo prima di tutto trovare una successione da associare a x^*. Per questo usiamo ancora la successione $e^{(n)}$ definita in (2.31) e definiamo

$$y_i = \overline{\langle\langle x^*, e^{(i)} \rangle\rangle}\,.$$

In questo modo si costruisce una successione $y = (y_i)$.

Proviamo prima di tutto che $y \in l^{p'}$, $||y|| \leq ||x^*||$. Proveremo poi che

$$\langle\langle x^*, x \rangle\rangle = \sum_{i=1}^{+\infty} \bar{y}_i x_i\,. \tag{2.35}$$

Introduciamo la successione $x^{(n)} \in l^p$ definita da

$$x^{(n)} = x_i^{(n)} \qquad \text{con} \quad x_i^{(n)} = \begin{cases} |y_i|^{p'-1} \frac{y_i}{|y_i|} & \text{se} \quad i \leq n \text{ e } y_i \neq 0 \\ 0 & \quad\text{altrimenti.} \end{cases}$$

Ovviamente, $x^{(n)} \in l^p$ per ogni n e

$$\langle\langle x^*, x^{(n)} \rangle\rangle \leq ||x^*|| \cdot ||x^{(n)}||_p$$

e, d'altra parte,

$$\langle\langle x^*, x^{(n)} \rangle\rangle = \sum_{i=1}^{n} |y_i|^{p'}\,.$$

Passando al limite rispetto ad n si trova

$$||y||_{p'} \le ||x^*||\,.$$

Dunque, $y \in l^{p'}$ e $||y|| \le ||x^*||$.

Proviamo ora che vale la (2.35). Fissato l'elemento $x \in l^p$, consideriamo la successione $x^{(n)}$ di elementi di l^p,

$$x^{(n)} = \sum_{i=0}^{n} x_i e^i\,, \quad \text{ossia} \quad x_r^{(n)} = \begin{cases} x_r & \text{se} \quad r \le n \\ 0 & \text{se} \quad r > n\,. \end{cases}$$

Si vede che, se $p < +\infty$,

$$x = \lim_{r \to +\infty} x^{(n)}\,.$$

Questo limite si calcola nella norma di l^p. Dunque, essendo x^* continuo,

$$\lim_n \langle\langle x^*, x^{(n)} \rangle\rangle = \langle\langle x^*, x \rangle\rangle\,.$$

Inoltre,

$$\lim_n \langle\langle x^*, x^{(n)} \rangle\rangle = \lim \sum_{r=0}^{+\infty} \bar{y}_r x_n^{(r)} = \lim \sum_{r=0}^{n} \bar{y}_r x_n = \sum_{i=0}^{+\infty} \bar{y}_r x_r$$

perchè si è già provato che $y \in l^{p'}$ e quindi che $x \to \sum_{i=0}^{+\infty} \bar{y}_r x_r$ è un funzionale continuo. Si trova così che vale la rappresentazione

$$\langle\langle x^*, x \rangle\rangle = \sum_{i=0}^{+\infty} \bar{y}_r x_r\,.$$

Osservazione 93 Nel caso particolare $p = 2$, una realizzazione del duale di l^2 è lo spazio stesso. ∎

2.7.0.2 Il duale di c_0

Ricordiamo che il simbolo c_0 indica il s.spazio di l^∞ i cui elementi sono le successioni che convergono a zero. Proviamo che il duale di c_0 è realizzato da l^1. Per provare questo notiamo che se $(\xi_k) \in l^1$ allora la trasformazione

$$x \to \sum_{k=1}^{+\infty} \bar{\xi}_k x_k \tag{2.36}$$

è lineare e continua su l^∞ e che la trasformazione da $\xi \in l^1$ al funzionale definito da (2.36) è isometrica e antilineare. Queste proprietà si conservano sostituendo l^∞ con c_0.

Dobbiamo provare che ogni $x^* \in (c_0)^*$ ammette la rappresentazione (2.36).

Notiamo prima di tutto che ogni l^p, $p < +\infty$, è un sottospazio di c_0 e che l'immersione di l^p in c_0 è continua. Dunque, ogni funzionale lineare continuo su c_0 è anche un funzionale lineare continuo su l^p, per ogni $p < +\infty$. Ciò suggerisce di porre ancora

$$\xi_i = \overline{\langle\langle x^*, e^{(i)}\rangle\rangle}\,.$$

Si trova così un vettore $\xi = (\xi_i)$, candidato ad essere un rappresentante di x^*.

Come si è detto, il vettore ξ è nel duale $l^{p'}$ di l^p per ogni $p < +\infty$. In particolare quindi è in l^∞. Proviamo che inoltre tale vettore è anche in l^1. Scegliamo per questo la seguente successione $x^{(n)}$ in c_0. L'elemento r-mo della successione $x^{(n)}$ è

$$x_r^{(n)} = \begin{cases} \frac{\xi_r}{|\xi_r|} & \text{se } r \le n \text{ e } \xi_r \neq 0 \\ 0 & \text{altrimenti} \end{cases} \qquad \text{e quindi} \quad x^{(n)} = \sum_{r\le n\ \ \xi_r\neq 0} \frac{\xi_r}{|\xi_r|} e^{(r)}\,.$$

Dato che $(\xi_n) \in l^\infty$, la successione $(x^{(n)})$ è una successione limitata in c_0. Esiste quindi un numero M tale che

$$\left|\langle\langle x^*, x^{(n)}\rangle\rangle\right| = \sum |\xi_r| < M$$

per ogni n. Ciò prova che $(\xi_r) \in l^1$.

Ora, per ogni $x \in c_0$, $x = (x_i)$, vale:

$$\langle\langle x^*, x\rangle\rangle = \lim_N \langle\langle x^*, \sum_{i=0}^N x_i e_i\rangle\rangle = \lim_N \sum_{i=0}^N \bar{\xi}_i x_i = \sum_{i=0}^{+\infty} \bar{\xi}_i x_i\,.$$

Ciò completa la dimostrazione.

Osservazione 94 Notiamo nuovamente che nella dimostrazione si usa la densità in c_0 della successione $e^{(i)}$. Questa successione è densa in l^p per $1 \le p < +\infty$ e anche in c_0; *ma non in* l^∞. ∎

2.7.0.3 Il duale di $L^p(\Omega)$, $1 \le p < +\infty$

In questo caso X^* è isometricamente isomorfo a $L^{p'}(\Omega)$, con

$$p' = \frac{p}{p-1} \quad \text{se } p > 1; \qquad L^\infty(\Omega) \quad \text{se } p = 1.$$

Ciò vale con $\Omega \subseteq \mathbb{R}^n$, limitato o meno.

Accenniamo alla dimostrazione nel caso $p = 1$ e $\Omega = (a, b)$ e nel caso in cui il campo scalare è $\mathbb{R}$.

E' ovvio che per ogni $\xi \in L^\infty(a, b)$, il funzionale su $L^1(a, b)$ definito da

$$x \to \int_a^b \xi(s)x(s)\,\mathrm{d}s$$

è continuo, di norma minore o uguale a $||\xi||_\infty$ e in realtà si vede che vale l'uguaglianza. Per non interrompere il filo del discorso, mostreremo ciò alla fine della dimostrazione.

Viceversa, sia x^* un funzionale lineare e continuo su $L^1(a, b)$. Dobbiamo associargli una funzione $\xi(s) \in L^\infty(a, b)$ tale che

$$\langle\langle x^*, x\rangle\rangle = \int_a^b \xi(s)x(s)\,\mathrm{d}s\,.$$

Introduciamo la famiglia delle funzioni $\chi_t(s)$, una funzione per ogni $t \in (a, b)$,

$$\chi_t(s) = \begin{cases} 1 & \text{se} \quad a < t < s \\ 0 & \qquad \text{altrimenti} \end{cases} \tag{2.37}$$

e studiamo i valori che x^* assume su queste funzioni. La ragione di ciò è che *le funzioni a costanti a tratti sono dense in* $L^1(a, b)$.

Associamo ad $x^* \in (L^1(a, b))^*$ la funzione

$$g(t) = \langle\langle x^*, \chi_t\rangle\rangle$$

E':

$$|g(t) - g(t')| = |\langle\langle x^*, \chi_t - \chi_{t'}\rangle\rangle \le ||x^*|| \cdot ||\chi_t - \chi_{t'}||_1 = ||x^*|| \cdot |t - t'|\,.$$

Dunque la funzione $g(t)$ è lipschitziana e quindi è assolutamente continua. Per essa vale

$$g(t) = \int_a^t \bar{\xi}(s)\,\mathrm{d}s$$

e inoltre $\bar{\xi}(s)$ è q.o. la derivata di $g(t)$. Dunque $\xi \in L^\infty(a, b)$ perché il rapporto incrementale di g è limitato.

Notiamo che

$$\langle\langle x^*, \chi_t\rangle\rangle = g(t) = \int_a^t \xi(s)\,\mathrm{d}s = \int_a^b \xi(s)\chi_t(s)\,\mathrm{d}s\,.$$

Una qualunque funzione a scala si rappresenta come combinazione lineare di funzioni χ_t:

$$\psi(t)=\sum_{i=1}^n \psi_i\chi_{t_i}(s)$$

e quindi

$$\langle\langle x^*,\psi\rangle\rangle=\sum_{i=1}^n \psi_i\int_a^b \xi(s)\chi_{t_i}(s)\,\mathrm{d}s=\int_a^b \xi(s)\psi(s)\,\mathrm{d}s\,.$$

Abbiamo già notato che il funzionale

$$x\to\int_a^b \xi(s)x(s)\,\mathrm{d}s$$

è continuo.

Sia ora $x\in L^1(a,b)$. Esiste una successione di funzioni a scala ψ_n convergente ad x in $L^1(a,b)$. Allora,

$$\langle\langle x^*,x\rangle\rangle=\lim_N\langle\langle x^*,\psi_N\rangle\rangle=\lim_N\int_a^b \xi(s)\psi_N(s)\,\mathrm{d}s=\int_a^b \xi(s)x(s)\,\mathrm{d}s\,.$$

Ciò è quanto volevamo provare.

Per completare la dimostrazione, mostriamo che la norma del funzionale

$$x\to\int_a^b \xi(s)x(s)\,\mathrm{d}s\qquad x\in L^1(a,b)$$

è uguale a $\|\xi\|_\infty$: basta provare che esiste una successione (x_n) in $L^1(a,b)$ tale che

$$\|x_n\|_{L^1(a,b)}=1\quad\text{e inoltre}\quad\lim\int_a^b \xi(s)x_n(s)\,\mathrm{d}s=\|\xi\|_\infty\,.$$

Ricordiamo che gli elementi di $L^\infty(a,b)$ sono classi di equivalenza, ma l'integrale è indipendente dal rappresentante della classe e quindi col simbolo ξ indichiamo sia la classe che uno qualsiasi dei suoi elementi.

Se $\xi=0$ il funzionale è nullo e quindi ha norma zero, uguale a $\|\xi\|_\infty$. Consideriamo quindi il caso $\|\xi\|_\infty>0$ e fissiamo un (qualsiasi rappresentante ξ della classe di equivalenza. Per ogni n esiste un insieme misurabile A_n tale che

$$\|\xi\|_\infty-\frac{1}{n}<|\xi(s)|\le\|\xi\|_\infty\,,\qquad \mu(A_n)>0\,.$$

Consideriamo la funzione

$$x_n(s)=\begin{cases}\frac{\operatorname{sgn}\xi(s)}{\mu(A_n)} & \text{se}\quad s\in A_n\\ 0 & \text{se}\quad s\notin A_n\,.\end{cases}$$

Si noti che
$$\|x_n\|_{L^1(a,b)} = 1 \qquad \text{per ogni } n.$$
Si ha:
$$\int_a^b \xi(s)x_n(s)\,\mathrm{d}s = \frac{1}{\mu(A_n)}\int_{A_n} |\xi(s)|\,\mathrm{d}s$$
e quindi
$$\|\xi\|_\infty - \tfrac{1}{n} = \tfrac{1}{\mu(A_n)}\int_{A_n} (\|\xi\|_\infty - 1/n)\,\mathrm{d}s \le \tfrac{1}{\mu(A_n)}\int_{A_n} |\xi(s)|\,\mathrm{d}s =$$
$$= \int_a^b \xi(s)x_n(s)\,\mathrm{d}s \le \|\xi\|_\infty \|x_n\|_{L^1(a,b)} = \|\xi\|_\infty .$$
Queste disuguaglianze mostrano che
$$\lim_{n\to+\infty} \int_a^b \xi(s)x_n(s)\,\mathrm{d}s = \|\xi\|_\infty$$
e ciò è quanto volevamo provare.

Osservazione 95 Una dimostrazione del tutto analoga porta ad identificare il duale di $L^p(\Omega)$, per ogni $p < +\infty$. Invece gli argomenti precedenti *non* si estendono al caso $L^\infty(\Omega)$ perché *le funzioni costanti a tratti non sono dense in* $L^\infty(\Omega)$ nemmeno si $\Omega = [a,b]$. ∎

2.7.0.4 Il duale di $C(a,b)$

La costruzione di un rappresentante per il duale di $C(a,b)$ è più profonda delle costruzioni precedenti, e richiede la conoscenza del *teorema di Hahn-Banach*, si veda il Teorema 118: *Sia X uno s.l.n. e sia Y un suo s.spazio. Sia L_0 un funzionale* lineare continuo *su Y. Esiste un'estensione di L ad X tale che*
$$||L||_{X^*} = \sup\{|Ly| \mid y \in Y\,,\ ||y||_X = 1\}\,.$$

Introduciamo ora una realizzazione del duale di $C(a,b)$, lo spazio di Banach delle funzioni continue sull'intervallo limitato e chiuso $[a,b]$, con la norma dell'estremo superiore. Fissiamo un elemento x^* del duale. Per rappresentarlo, procediamo in tre passi:

PASSO 1) Introduciamo lo spazio lineare di tutte le funzioni limitate su $[a,b]$, continue o meno, dotato della norma dell'estremo superiore. Si trova uno s.l.n. che indicheremo col generico simbolo B.

$C(a,b)$ essendo un s.spazio di B, il funzionale lineare e continuo x^*, definito su $C([a,b])$, si estende ad un funzionale lineare continuo su B, con la stessa norma, per il Teorema di Hahn-Banach. Tale estensione non è unica. Fissiamone una, che indichiamo col simbolo $\tilde{x}^*$.

Per ogni $t \in [a,b]$, introduciamo le funzioni definite come in (2.37) e la funzione

$$v(t) = \langle\langle \tilde{x}^*, \chi_t \rangle\rangle \,. \tag{2.38}$$

Sia ora $f \in C(a,b)$. Essendo $[a,b]$ compatto, la funzione f è uniformemente continua e quindi si approssima in modo uniforme con funzioni costanti a tratti. Queste possono costruirsi scegliendo un insieme finito $\{t_i\}_{i=1}^n$ di punti di $[a,b]$, abbastanza fitto, e quindi definendo

$$z_n(t) = f(t_{i-1}) \qquad \forall t \in [t_{i-1}, t_i)$$

ossia

$$z_n(t) = \sum_{i=1}^n f(t_{i-1})[\chi_{t_i}(s) - \chi_{t_{i-1}}(s)] \,.$$

Si ha quindi

$$\langle\langle x^*, f \rangle\rangle = \langle\langle \tilde{x}^*, f \rangle\rangle = \lim \langle\langle \tilde{x}^*, z_n \rangle\rangle = \lim \sum_{i=1}^n f(t_{i-1})[\langle\langle \tilde{x}^*, \chi_{t_i}(s) \rangle\rangle - \langle\langle \tilde{x}^*, \chi_{t_{i-1}}(s) \rangle\rangle]$$

$$= \lim \sum_{i=1}^n f(t_{i-1})[v(t_i) - v(t_{i-1})] \,.$$

Si noti che nel caso particolare in cui $v(t) = t$, tale limite è $\int_a^b f(s)\,\mathrm{d}s$.

PASSO 2) Introduciamo un simbolo per indicare il limite precedente,

$$\int_a^b f \,\mathrm{d}v = \lim \sum_{i=1}^n f(t_{i-1})[v(t_i) - v(t_{i-1})] \tag{2.39}$$

(ovviamente, il limite non dipende dalla partizione scelta per definirlo, dato che esso deve essere $\langle\langle x^*, f \rangle\rangle$).

Il particolare integrale definito da (2.39) si chiama INTEGRALE DI STILTJES.

Si osservi che la rappresentazione di x^* come integrale di Stiltjes usa la continuità uniforme di f; e quindi in generale $\tilde{x}^*$ non avrà tale rappresentazione.

PASSO 3) Ricapitolando, abbiamo rappresentato ogni elemento del duale di $C(a,b)$ come un integrale di Stiltjes. Dobbiamo ora studiare le proprietà di tale integrale, per trovare uno spazio di Banach che realizzi $[C(a,b)]^*$.

Per la funzione $v(t)$ definita in (2.38) si ha

$$\sum_{i=1}^{n} |v(t_i) - v(t_{i-1})| = \sum_{i=1}^{n} \{\operatorname{sgn}[v(t_i) - v(t_{i-1})]\}[v(t_i) - v(t_{i-1})]$$
$$\sum_{i=1}^{n} \langle\langle \tilde{x}^*, \left\{\operatorname{sgn}[v(t_i) - v(t_{i-1})][\chi_{t_i} - \chi_{t_{i-1}}]\right\}\rangle\rangle$$
$$= \langle\langle \tilde{x}^*, \sum_{i=1}^{n} \left\{\operatorname{sgn}[v(t_i) - v(t_{i-1})][\chi_{t_i} - \chi_{t_{i-1}}]\right\}\rangle\rangle$$
$$\leq ||\tilde{x}^*|| \sup_s \sum_{i=1}^{n} |\chi_{t_i}(s) - \chi_{t_{i-1}}(s)| = ||\tilde{x}^*|| = ||x^*||$$

perché la differenza $|\chi_{t_i}(s) - \chi_{t_{i-1}}(s)|$ vale 1 oppure 0.

Questa disuguaglianza vale *per ogni* suddivisione dell'intervallo $[a,b]$ in un numero finito di punti e quindi esiste un numero M, $M = ||x^*||$, tale che

$$V_a^b v = \sup_{\{t_i\}} \sum |v(t_i) - v(t_{i-1})| < M\,.$$

Introduciamo ora una definizione che si applica ad una qualsiasi funzione definita su $[a,b]$, non necessariamante alla funzione definita in (2.38).

Definizione 96 Sia $f(t)$ una funzione definita su $[a,b]$. La VARIAZIONE TOTALE di f su $[a,b]$ è

$$V_a^b f = \sup_{\{t_i\}} \sum |f(t_i) - f(t_{i-1})| \leq +\infty\,.$$

L'estremo superiore è calcolato rispetto a tutte le sequenze finite $a \leq t_0 < t_1 < t_n < \cdots < t_n \leq b$ di punti dell'intervallo $[a,b]$.

Se $V_a^b f < +\infty$ la funzione f si dice FUNZIONE A VARIAZIONE LIMITATA SU $[a,b]$.

La struttura delle funzioni a variazione limitata è stata studiata con estrema precisione. Si prova in particolare che ogni funzione a variazione limitata è differenza di due funzioni monotone e che, quindi, i suoi punti di discontinuità sono salti. Si prova inoltre che

$$\sup_{||f||<1} \int_a^b f\,\mathrm{d}v = V_a^b v$$

e questo suggerisce di scegliere come spazio per rappresentare $[C(a,b)]^*$ uno spazio di funzioni a variazione limitata. Infatti abbiamo provato che

$$\langle\langle x^*, f\rangle\rangle = \int_a^b f \, \mathrm{d}v$$

ove v è la funzione definita in (2.38). Bisogna però notare che può aversi

$$\int_a^b f \, \mathrm{d}v = \int_a^b f \, \mathrm{d}v' \qquad \forall f \in C(a,b)$$

anche con $v \neq v'$. E quindi la rappresentazione che abbiamo trovato per x^* non è unica. Si prova però che l'uguaglianza può aversi, per ogni f, solo se v e v' differiscono per il valore che assumono in un punto di salto oppure nell'estremo sinistro a dell'intervallo. Ciò suggerisce di definire

$$NV(a,b)$$

lo spazio delle FUNZIONI A VARIAZIONE LIMITATA NORMALIZZATE su $[a,b]$, ossia continue a sinistra e nulle in a, dotato della norma

$$V_a^b(v)\,.$$

Si prova che questo spazio è di Banach e che vale:

Teorema 97 (di Riesz) *Lo spazio $NV(a,b)$ è una realizzazione del duale di $[C(a,b)]^*$ e ogni $x^* \in [C(a,b)]^*$ si rappresenta (in modo unico) come*

$$\langle\langle x^*, f\rangle\rangle = \int_a^b f \, dv\,, \qquad v \in NV(a,b)\,.$$

2.7.0.5 Il duale di $C(K)$

Ricordiamo che quando si usa il simbolo $C(K)$ si intende che l'insieme K è compatto.

Non abbiamo gli strumenti per studiare il duale di $C(K)$. Possiamo però descrivere come si rappresenta l'azione su $C(K)$ di un elemento x^* del suo duale. Per ogni $x^* \in (C(K))'$ si trovano una misura di Borel m ed una funzione $\psi(s)$ misurabile secondo Borel su K e tale che

$$|\psi(s)| = 1 \qquad \text{q.o. } s \in K$$

e per la quale vale

$$\langle\langle x^*, x\rangle\rangle = \int_K \psi(s)x(s) \, \mathrm{d}m$$

2.8 Il teorema di Baire e le sue conseguenze

Una semplice osservazione che vale in $\mathbb{R}^2$ è la seguente: gli iperpiani per 0 in questo caso sono rette di equazione $y = mx$ oppure $x = 0$. Esse sono parametrizzate dal punto in cui intersecano la circonferenza $x^2 + y^2 = 1$. Dunque $\mathbb{R}^2$ non è unione di una famiglia numerabile di rette per 0; e questa osservazione si generalizza a rette qualsiasi, ed a dimensione $n > 2$. Vediamo come questo risultato si estende ad un generico spazio di Banach.

Proveremo il teorema seguente, non ovvio nemmeno in dimensione finita:

Teorema 98 (di Baire) *Sia X uno spazio di Banach e sia (A_n) una successione di s.insiemi di X, ciascuno dei quali è chiuso e privo di punti interni. Allora,*

$$\bigcup A_n \neq X\,.$$

Rimandando alla fine di questo paragrafo la dimostrazione, illustriamo varie conseguenze importanti di questo teorema.

Notiamo prima di tutto che un s.spazio di X, diverso da X stesso, non ha punti interni. Dunque vale in un generico spazio di Banach la proprietà che abbiamo notato sopra per $\mathbb{R}^2$, che conviene enunciare come segue:

Teorema 99 *Sia (X_n) una successione di s.spazi di uno spazio di Banach X. Se $X = \cup X_n$ allora esiste n_0 tale che $X = X_{n_0}$.*

Il Teorema di Baire è un potente strumento per lo studio delle proprietà degli operatori lineari tra due spazi di Banach X ed Y. Esso talvolta si usa direttamente; più spesso interviene grazie ai quattro teoremi seguenti. Il primo che presentiamo va sotto il nome di Teorema di Banach-Steinhaus. Esso concerne s.insiemi $\mathcal{A}$ di $\mathcal{L}(X,Y)$. Ricordiamo che $\mathcal{L}(X,Y)$ è uno spazio normato e quindi ha senso investigare quando $\mathcal{A}$ è un s.insieme limitato di $\mathcal{L}(X,Y)$. Ciò avviene se esiste M tale che $||A|| \leq M$ per ogni $A \in \mathcal{A}$.

Fissiamo ora un qualsiasi elemento $x \in X$ e consideriamo l'insieme dei "valori" Ax, $A \in \mathcal{A}$. Questo è un s.insieme di Y che è limitato se l'insieme $\mathcal{A}$ è limitato in $\mathcal{L}(X,Y)$. Infatti, $||Ax||_Y \leq ||A|| \cdot ||x|| \leq M \cdot ||x||$ per ogni $A \in \mathcal{A}$. Il teorema di Banach-Steinhaus permette di invertire questa proprietà:

Teorema 100 (di Banach-Steinhaus) *Sia $\mathcal{A}$ un s.insieme di $\mathcal{L}(X,Y)$. Supponiamo che per ogni $x \in X$ esista un numero M_x tale che*

$$||Ax|| \leq M_x \qquad \forall A \in \mathcal{A}\,. \tag{2.40}$$

(Sottolineiamo: M_x indipendente *da $A \in \mathcal{A}$). In questo caso $\mathcal{A}$ è un s.insieme limitato di $\mathcal{L}(X,Y)$.*

Dim. Indichiamo con $X_n \subseteq X$ l'insieme

$$X_n = \{x \mid ||Ax|| \leq n \quad \forall A \in \mathcal{A}\} .$$

La condizione (2.40) mostra che

$$\bigcup X_n = X .$$

Consideriamo ora un operatore $A \in \mathcal{A}$. Essendo A continuo, l'insieme $\{x \mid ||Ax|| \leq n\}$ è chiuso e quindi

$$X_n = \bigcap_{A \in \mathcal{A}} \{x \mid ||Ax|| \leq n\}$$

è esso stesso chiuso. Abbiamo quindi una famiglia di chiusi la cui unione è X. Per il *Teorema di Baire,* uno almeno deve avere punti interni. Sia esso X_N. Esiste $x_0 \in X_N$ ed esiste $\epsilon > 0$ per cui

$$\|x\| < \epsilon \quad \Longrightarrow \quad x_0 + x \in X_N .$$

Dunque, se $||x|| < \epsilon$ si ha

$$||Ax|| \leq ||A(x + x_0)|| + ||Ax_0|| \leq N + ||Ax_0|| = N + M_{x_0} = M ,$$

con M *indipendente* da A. Ciò prova la limitatezza del s.insieme $\mathcal{A}$ di $\mathcal{L}(X, Y)$. ∎

Il Teorema di Banach-Steinhaus permette di passare da un'informazione *puntuale,* la limitatezza dell'insieme dei valori assunti in ciascun punto x, ad una limitatezza *uniforme* sulla sfera $\{x \mid ||x|| \leq 1\}$. Per questo esso va anche sotto il nome di Teorema della limitatezza uniforme.

Osservazione 101 (Interpretazione geometrica) Il teorema di Banach-Steinhaus si può interpretare geometricamente come segue. Immaginiamo di riportare i valori $\|Ax\|$ su una semiretta verticale, e quindi immaginiamo che lo spazio X sia messo in orizzontale. Sopra ogni x si trova l'insieme dei valori $\{\|Ax\| ,\ A \in \mathcal{A}\}$ e quest'insieme è contenuto in un segmento di altezza M_x, variabile al variare di x. Per ipotesi, ciascuno di questi segmenti è limitato, $M_x < +\infty$.

Il teorema di Banach-Steinhaus asserisce che essi sono *uniformemente limitati,* ossia che sono più corti di un medesimo numero M. ∎

In dimensione finita una trasformazione lineare invertibile non può "schiacciare" un aperto trasformandolo in un s.insieme di un s.spazio proprio. Si ricordi il ruolo importante di questa proprietà nella dimostrazione del teorema della funzione inversa e della funzione implicita.

Una proprietà analoga vale anche in spazi di Banach:

Teorema 102 (della MAPPA APERTA) *Siano X ed Y spazi di Banach e sia $A \in \mathcal{L}(X,Y)$. Se A è suriettiva allora l'immagine di ogni aperto di X è un aperto di Y.*

Posponiamo la dimostrazione presentando invece due conseguenze del Teorema di Baire che si provano più facilmente mediante il teorema della Mappa aperta. Esse riguardano questo problema: abbiamo visto che gli operatori lineari tra X ed Y possono essere discontinui se X ha dimensione infinita. Gli esempi che abbiamo visto di operatori lineari discontinui sono però esempi di operatori il cui dominio non è tutto X. Supponiamo che il dominio dell'operatore sia tutto lo spazio. Ci chiediamo se in tal caso l'operatore (lineare) debba essere continuo. La risposta è negativa. Infatti si ha:

Teorema 103 *Siano X ed Y spazi di Banach. Esistono operatori lineari da X in Y, definiti su X e* non continui.

Si veda l'osservazione 141.
Però:

Teorema 104 (di BANACH) *Siano X ed Y spazi di Banach e sia $A \in \mathcal{L}(X,Y)$ una trasformazione lineare* iniettiva *da X in Y. Se l'immagine di A è* chiusa *allora la trasformazione lineare A^{-1} (definita su* im$A$*) è* continua.

<u>Dim</u>. Ricordiamo che, per definizione di $\mathcal{L}(X,Y)$, *un operatore $A \in \mathcal{L}(X,Y)$ ha dominio* uguale *ad X.*

L'immagine di una trasformazione lineare A è un s.spazio che in generale non è chiuso. *Il teorema assume invece che l'immagine di A sia un un s.spazio chiuso.* Dunque l'immagine di A è essa stessa uno spazio di Banach. Sostituendo Y con im$A$, possiamo supporre che A sia anche suriettiva.

La *trasformazione inversa* di A^{-1}, che è A, è suriettiva: per il teorema della mappa aperta, $A = (A^{-1})^{-1} \in \mathcal{L}(X,Y)$ trasforma aperti in aperti; e quindi A^{-1} è continua. ∎

Diamo ora un test importante per provare la continuità direttamente di un operatore (e non del suo operatore inverso). Premettiamo quest'osservazione:

Teorema 105 *Siano X ed Y spazi di Banach e sia A: $X \mapsto Y$. Allora:*

1. *se $A \in \mathcal{L}(X,Y)$ (e quindi* dom$A = X$*) allora il* grafico *di A è chiuso in $X \times Y$.*

2. *se A: $X \mapsto Y$ ha grafico chiuso ed inoltre è continuo allora il suo dominio è un insieme chiuso.*

Dim. Il primo asserto è ovvio e si lascia per esercizio. Proviamo il secondo. Sia (x_n) una successione in dom$A$ che è convergente, $x_n \to x_0$ e sia $y_n = Ax_n$. La continuità di A mostra che (y_n) è convergente, $y_n \to y_0$. Dunque $(\,(x_n, y_n)\,)$ è una successione convergente in $\mathcal{G}(A)$, che è chiuso. E quindi $(x_0, y_0) \in \mathcal{G}(A)$. In particolare $x_0 \in$ dom$A$.

Si noti che la linearità di A non ha un ruolo in questa seconda parte della dimostrazione. ∎

Viceversa, diamo ora una condizione che permette di concludere la continuità di un operatore il cui grafico è chiuso:

Teorema 106 (del GRAFICO CHIUSO) *Siano X ed Y spazi di Banach e sia A un operatore lineare da X in Y,* con dominio uguale ad X. *Se il grafico di A è chiuso allora A è continuo.*

Dim. Indichiamo con G il grafico di A. Per ipotesi, G è un s.spazio chiuso dello spazio di Banach $X \times Y$; e quindi è esso stesso uno spazio di Banach.

Introduciamo i due operatori, ovviamente lineari e continui:

$$P\ : G \to Y\,, \qquad P(x, Ax) = Ax$$
$$\Pi\ : G \to X\,, \qquad \Pi(x, Ax) = x\,.$$

Oltre che continuo, l'operatore Π è *suriettivo,* perché dom$A = X$ per ipotesi; ed è iniettivo perchè se $Ax_1 \neq Ax_2$ allora $x_1 \neq x_2$. Dunque esiste Π^{-1} e, per il Teorema di Banach, Π^{-1} è continuo. Dunque,

$$Ax = P(\Pi^{-1}x)$$

è continua. ∎

L'esempio 108 seguente mostra che esistono operatori lineari *il cui grafico è chiuso* ma che *non sono continui.* Definiamo quindi:

Definizione 107 *Sia A uno operatore lineare tra due spazi di Banach X ed Y. L'operaore A si dice* CHIUSO *quando il suo grafico è chiuso in $X \times Y$.* ∎

E ora mostriamo che operatori chiusi ma non continui non solo esistono ma sono anche importanti per le applicazioni. Il Teorema 106 mostra che il loro dominio non può essere chiuso.

Esempio 108 Sia $X = Y = C(0,1)$ e sia

$$\text{dom}A = C^1(0,1)\,, \qquad Ax = x'\,.$$

Se $x \in \mathrm{dom} A$, vale

$$x(t) = x(0) + \int_0^t x'(s)\,\mathrm{d}s\,.$$

Come si verifica facilmente, l'operatore A non è continuo. Proviamo che il suo grafico è chiuso. Consideriamo quindi una successione nel grafico che è convergente e mostriamo che essa converge ad un punto del grafico. Sia quindi

$$(x_n, Ax_n) \to (x_0, y_0)\,.$$

Dobbiamo provare che $x_0 \in \mathrm{dom} A$ e che $Ax_0 = y_0$.

Si noti che y_0 è limite uniforme delle funzioni continue Ax_n e quindi y_0 è una funzione continua. Dunque dobbiamo provare che si può scrivere

$$x_0(t) = x_0(0) + \int_0^t y_0(s)\,\mathrm{d}s\,.$$

Questa uguaglianza segue da

$$x_n(t) = x_n(0) + \int_0^t x_n'(s)\,\mathrm{d}s$$

e da

$$\begin{cases} x_n \to x_0 & \text{uniformemente su } [0,1] \\ x_n' \to y_0 & \text{uniformemente su } [0,1]. \end{cases} \blacksquare$$

Osserviamo infine:

Corollario 109 *Siano X, Y e Z tre spazi di Banach. Sia $A \in \mathcal{L}(X, Y)$ e sia B un operatore lineare chiuso da Y in Z. Se*

$$\mathrm{im}\, A \subseteq \mathrm{dom}\, B$$

allora l'operatore composto BA è continuo.

Dim. Si vede facilmente che BA è definito su X, ed è chiuso. Dunque è continuo. ∎

2.8.1 Proiezioni

Un operatore $P \in \mathcal{L}(X)$ si dice una PROIEZIONE se

$$P^2 = P\,.$$

Si noti che le proiezioni vengono sempre a coppie. Infatti,

Teorema 110 *L'operatore P è una proiezione se e solo se l'operatore $I-P$ è una proiezione. Inoltre,* $\mathrm{im}P \cap \mathrm{im}(I-P) = \{0\}$.

<u>Dim</u>. Infatti,

$$(I-P)(I-P) = I - P - P + P^2 = I - P$$

se e solo se $P^2 = P$ ossia se e solo se P è una proiezione.

Se $x = Px' = (I-P)x''$ allora

$$Px' = x'' - Px'' \qquad \text{da cui} \quad Px' = Px'' - P^2x'' = Px'' - Px'' = 0$$

e quindi $x = Px' = 0$. ∎

Inoltre:

Teorema 111 *L'immagine di una proiezione è un s.spazio chiuso di X.*

<u>Dim</u>. Sia infatti (Px_n) una successione in $\mathrm{im}P$, $Px_n \to y_0$. Dobbiamo provare che $y_0 \in \mathrm{im}P$.

Poniamo $y_n = Px_n$ e notiamo che

$$Py_n = P^2x_n = Px_n \to y_0$$

e d'altra parte, essendo P continua, $Py_n = P^2y_n \to Py_0$. Dunque, $y_0 = Py_0 \in \mathrm{im}P$. ∎

Di conseguenza, *ogni proiezione identifica sempre una coppia di s.spazi chiusi: l'immagine di P e quella di $(I-P)$.* Questi s.spazi hanno in comune solo l'elemento 0. Inoltre, ogni x si rappresenta come

$$x = Px + (I-P)x\,.$$

Questa formula suggerisce un legame tra operatori di proiezione e complementare. Vale infatti:

Teorema 112 *L'immagine di una proiezione P è un s.spazio chiuso di X, dotato di complementare chiuso. Viceversa, sia X_1 un s.spazio di X chiuso e dotato di complementare chiuso X_2. Esiste una proiezione P la cui immagine è X_1.*

<u>Dim</u>. Sia P una proiezione ed $X_1 = \mathrm{im}P$. Definiamo

$$X_2 = \mathrm{im}(I-P)\,.$$

Si è già visto che $X_1 \cap X_2 = \{0\}$ e

$$X_1 + X_2 = \{Px + (I - P)y \mid x \in X\ ,\ y \in Y\} = X\ .$$

Dunque, l'immagine di P è un s.spazio dotato di complementare.

Viceversa, sia

$$X = X_1 \oplus X_2\ ,$$

somma diretta di due s.spazi chiusi. Questo vuol dire che per ogni x esistono x_1 ed x_2 unici e tali che

$$x = x_1 + x_2\ . \tag{2.41}$$

Si definisca $Px = x_1 = x_1 + 0$. Segue da qui che $P(Px) = x_1$.

L'operatore P è lineare perchè se $x' = x_1' + x_2'$, $x'' = x_1'' + x_2''$, allora $\alpha x' + \beta x'' = \alpha(x_1' + x_2') + \beta(x_1'' + x_2'') = (\alpha x_1' + \beta x_2') + (\alpha x_1'' + \beta x_2'')$; e quindi $P(\alpha x' + \beta x'') = (\alpha x_1' + \beta x_2') = \alpha P x' + \beta P x''$. Se possiamo provare la continuità di P, abbiamo che P è una proiezione.

L'operatore P è definito su X e quindi, per provare che è continuo, basta provare che è chiuso. Sia quindi (x_n) una successione convergente ad x_0 e sia $y_n = Px_n$. Supponiamo che (y_n) converga ad y_0.

L'uguaglianza (2.41) mostra l'esistenza di un elemento $z_n \in X_2$ tale che

$$x_n = Px_n + z_n = y_n + z_n\ .$$

Di conseguenza, anche la successione (z_n) converge, a $z_0 = x_0 - y_0$.

I s.spazi essendo chiusi, vale $y_0 \in X_1$, $z_0 \in X_2$. Essendo inoltre

$$x_0 = y_0 + z_0\ , \qquad \text{si ha} \quad Px_0 = y_0\ .$$

Segue da qui che l'operatore P è chiuso e quindi continuo; dunque è una proiezione. ∎

Ne segue:

Corollario 113 *Se P è una proiezione allora $X = \mathrm{im}P$ ammette complementare e questo è $\mathrm{im}(I - P)$:*

$$X = \mathrm{im}P \oplus \mathrm{im}(I - P)\ .$$

E' importante osservare che in generale niente può dirsi della norma di una proiezione. L'esempio seguente mostra che esistono proiezioni di norma arbitrariamente grande.

Esempio 114 Sia $X = \mathbb{R}^2$ normato dalla usuale norma

$$||(\xi, \eta)|| = \sqrt{\xi^2 + \eta^2}$$

e siano

$$X_1 = \{(\xi, 0) \mid \xi \in \mathbb{R}\} , \qquad X_2 = \{r(\cos\theta, \sin\theta) \mid r \in \mathbb{R}\}$$

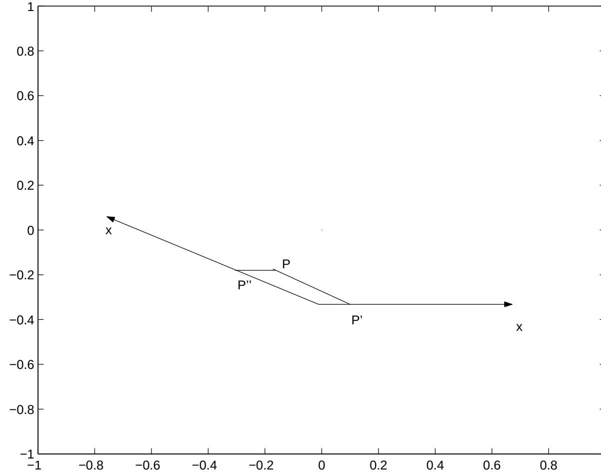
ove $\theta \in (0, \pi/2)$ è fissato. Dunque, X_2 è una retta per l'origine, non coincidente con X_1.

Ogni punto $x = (\xi, \eta)$ può rappresentarsi nella forma

$$x = (\xi - \frac{\eta}{\sin\theta}\cos\theta, 0) + \frac{\eta}{\sin\theta}(\cos\theta, \sin\theta) ,$$

si veda la figura seguente.

Figura 2.3:



L'operatore P:

$$P(\xi, \eta) = (\xi - \frac{\eta}{\sin\theta}\cos\theta, 0)$$

è una proiezione.

L'operatore P dipende dalla scelta di θ, $P = P_\theta$.

La norma di P_θ è

$$\max_{\xi^2+\eta^2=1} \frac{\left|\xi - \frac{\eta}{\sin\theta}\cos\theta\right|}{\sqrt{\xi^2 + \eta^2}} \geq |\operatorname{cotg}\theta| .$$

Dunque,

$$\lim_{\theta\to 0} ||P_\theta|| = +\infty . \quad \blacksquare$$

2.8.2 Appendice: Applicazioni

Il Teorema di Baire e le sue conseguenze sono strumenti potenti per provare l'esistenza di oggetti dalle proprietà "strane". Mostriamo due esempi. Le dimostrazioni sono alquanto tecniche (e posposte) ma in questo caso piuttosto dei dettagli della dimostrazione è importante capirne l'idea.

Si *costruiscono* "esplicitamente", come somma di serie uniformemente convergenti di funzioni continue, delle funzioni che, pur essendo continue, non hanno derivata in nessun punto. Una dimostrazione, dovuta a Banach, dell'*esistenza* di tali funzioni si basa sul Teorema di Baire.

Teorema 115 *Esistono funzioni continue su un intervallo* $[a, b]$*, ovunque prive di derivata.*

Dim. (idea)

Ricordiamo che una funzione è derivabile nel punto x quando

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \quad \text{è un numero}$$

ossia quando il limite del rapporto incrementale esiste *finito.* Dunque, se $f(x)$ è derivabile in un punto x_0 allora si ha

$$|f(x_0 + h) - f(x_0)| < (1 + |f'(x_0)|)\,|h|$$

almeno per $|h|$ abbastanza piccolo.

Indichiamo con A_n l'insieme delle funzioni $f(x)$ continue su $[0, 1]$ con questa proprietà: $f \in A_n$ quando esiste un punto $x = x_f \in [0, 1)$ ed esiste $h > 0$ tali che

$$|f(x + h) - f(x)| < nh\,.$$

Una funzione derivabile in un punto (anche in uno solo) appartiene ad alcuni degli insiemi A_n: quello per cui $n > |f'(x)|$

Si prova che ciascuno degli insiemi A_n è chiuso.

Se ogni funzione continua è derivabile in almeno un punto allora $\cup_{n>0} A_n = C(0, 1)$ e quindi per il teorema di Baire deve esistere N tale che A_N ammette punti interni.

Notiamo ora che ogni funzione continua si approssima uniformemente, ossia nella norma di $C(0, 1)$, mediante funzioni "seghettate in modo fitto fitto" e quindi certamente non appartenenti ad A_N, se i "denti della sega" sono abbastanza ripidi.

Dunque nessun A_N ammette punti interni e quindi esiste in $C(0, 1)$ una funzione che non appartiene a nessuno degli insiemi A_N, e quindi continua ma priva di derivata in ciascun punto. ∎

La dimostrazione precedente usa in modo diretto il Teorema di Baire. Talvolta, anche per provare l'esistenza di esempi "strani" conviene usarne le conseguenza. La dimostrazione del teorema seguente si fa in modo più semplice usando il Teorema di Banach-Steinhaus.

Sia $f(x)$ una funzione continua su $[-\pi, \pi]$. Si associ ad essa la serie

$$\sum_{n=-\infty}^{+\infty} f_n e^{inx}\,, \qquad f_n = \frac{1}{2\pi}\int_{-\pi}^{+\pi} f(s)e^{-ins}\,\mathrm{d}s$$

che si chiama la SERIE DI FOURIER della funzione $f(x)$. Sotto ipotesi di regolarità, per esempio se la funzione $f(x)$ è di classe C^1 e inoltre $f(-\pi) = f(\pi)$, la serie converge uniformemente ad $f(x)$, nel senso che[15]

$$f(x) = \lim_{N\to+\infty} \sum_{n=-N}^{N} f_n e^{inx}\,.$$

La condizione che f sia di classe C^1 può indebolirsi, *ma non fino alla sola continuità.* Infatti:

Teorema 116 *Sia $x_0 \in [-\pi, \pi]$. Esiste una funzione f continua su $[-\pi, \pi]$ tale che $f(0) = f(2\pi)$ e tale che inoltre la serie di Fourier ad essa associata non converge in x_0.*

Dim. (idea) Lavoriamo nello spazio di Banach $C_P(-\pi, \pi)$ i cui elementi sono funzioni continue che verificano

$$f(-\pi) = f(\pi)$$

(la norma è la norma della convergenza uniforme. Si vede facilmente che questo è un s.spazio chiuso di $C(-\pi, \pi)$, e quindi esso stesso uno spazio di Banach).

Usando le formule di Eulero e la definizione di f_n, si vede che

$$\sum_{n=-N}^{N} f_n e^{inx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)\frac{\sin[(N+1/2)(x-s)]}{\sin[(x-s)/2]}\,\mathrm{d}s$$

e questo suggerisce di considerare gli operatori lineari e continui F_N da $C_P(-\pi, \pi)$ in sé definiti da:

$$(F_N f)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)\frac{\sin[(N+1/2)(x-s)]}{\sin[(x-s)/2]}\,\mathrm{d}s\,.$$

[15] è essenziale che la somma parziale sia construita con gli indici simmetrici rispetto all'indice zero.

Non è difficile vedere che ciascun operatore F_N è continuo su $C_P(-\pi, \pi)$.

Se per una funzione f il limite

$$\lim_{N \to +\infty} F_n f$$

esiste in $C_P(-\pi, \pi)$ allora esiste M_f tale che

$$\|F_N f\|_{C_P(-\pi,\pi)} < M_f \, .$$

Per il Teorema di Banach-Steinhaus, deve esistere M tale che

$$\|F_N\|_{C_P(-\pi,\pi)} < M \, .$$

Notando che

$$\lim_{x \to 0} \frac{\sin(N + 1/2)x}{\sin x/2} = 2N + 1$$

si intuisce che la famiglia di operatori $\mathcal{A} = \{F_N\}$ non è limitata. Quest'intuizione è corretta (anche se la verifica è macchinosa) e quindi esiste una funzione $f \in C_P(-\pi, \pi)$ che non è limite uniforme della sua serie di Fourier. ∎

2.8.3 Dimostrazioni posposte

Dimostrazione del TEOREMA 98, Teorema di Baire.

Premettiamo un'osservazione: Sia B_n una successione di palle contenuta ciascuna nella precedente:

$$B_n = \{x \mid ||x - x_n|| < \epsilon_n\} \subseteq \{x \mid ||x - x_{n-1}|| < \epsilon_{n-1}\} = B_{n-1}$$

così che

$$||x_n - x_{n+m}|| < \epsilon_n \, .$$

Sia $\lim \epsilon_n = 0$. Da

$$||x_n - x_{n+m}|| < \epsilon_n$$

si vede che la successione (x_n) è fondamentale ossia convergente,

$$\lim x_n = \tilde{x} \, .$$

Notiamo che il punto $\tilde{x}$ appartiene alla chiusura di B_n per ogni n e quindi si ha anche

$$\tilde{x} \in \bigcap_n \operatorname{cl} B_n \, .$$

Sia ora (A_n) una successione di insiemi chiusi e (B_n) una successione di palle con le proprietà appena dette e tali che, inoltre,

$$(\mathrm{cl}\, B_n) \cap A_n = \emptyset \,. \tag{2.42}$$

Allora, $\tilde{x}$ non appartiene a A_n per nessun n, grazie alla (2.42):

$$\tilde{x} \notin \bigcup A_n \,.$$

Infatti, se $\tilde{x} \in \cup A_n$ allora $\tilde{x} \in A_{n_0}$ per almeno un indice n_0. D'altra parte, si sa anche che $\tilde{x} \in \mathrm{cl}\, B_{n_0}$ e ciò non può essere perché gli insiemi A_{n_0} e $\mathrm{cl}\, B_{n_0}$ sono disgiunti.

Per provare il Teorema di Baire, costruiamo una successione di palle B_n che ha le proprietà dette sopra rispetto alla successione di insiemi (A_n), chiusi e privi di punti interni. Ciò porterà a trovare che $\tilde{x} \notin \cup A_n$ e quindi $\cup A_N \neq X$. Scegliamo $x_1 \notin A_1$ e una palla B_1 di centro x_1 e raggio minore di 1, tale che $(\mathrm{cl}\, B_1) \cap A_1 = \emptyset$. Sia $\epsilon_1 > 0$ il suo raggio. Non è restrittivo assumere $\epsilon_1 < 1$.

La palla B_1 esiste perché A_1 è chiuso e, essendo privo di punti interni, non è uguale ad X.

La palla B_1 non è contenuta in A_2 *perché* A_2 *non ha punti interni.* Dunque in B_1 esite un punto $x_2 \notin A_2$ e quindi interno al complementare dell'*insieme chiuso* A_2. Possiamo quindi scegliere una palla B_2 di centro x_2, contenuta in B_1 e di raggio minore di $\epsilon_1/2 < 1/2$, tale che $(\mathrm{cl}\, B_2) \cap A_2 = \emptyset$.

Sia $\epsilon_2 > 0$ il raggio di B_2.

Procedendo per induzione, scelti i punti $x_1, \ldots, x_k$ e le corrispondenti palle $B_1, \ldots, B_k$, scegliamo in B_k un punto $x_{k+1} \notin A_{k+1}$ e una sfera di centro x_{k+1} e raggio minore di $\epsilon_k/2 < 1/2^k$, tale che $(\mathrm{cl}\, B_{k+1}) \cap A_{k+1} = \emptyset$. Sia $\epsilon_{k+1} > 0$ il raggio di questa sfera.

La costruzione dei punti x_k e delle palle B_k può farsi perché gli A_k non hanno punti interni e sono chiusi.

Dato che $\epsilon_k \to 0$, esiste $\tilde{x} = \lim x_k$ e $\tilde{x} \in \cap_n \mathrm{cl}\, B_n$. Per quanto osservato sopra, $\tilde{x} \notin \cup A_n$, ossia

$$\bigcup A_n \neq X \,,$$

come si voleva provare. ∎

<u>Dimostrazione del TEOREMA 102</u>, Teorema della mappa aperta.

In questa dimostrazione interverrà la "differenza algebrica" di insiemi C, D di Y:

$$C - D = \{c - d \mid c \in C \,,\ d \in D\} = \bigcup_{d \in D} (C - d) \,.$$

Si noti che $C - C \neq \emptyset$ (e anche che $C - C \neq 0$, salvo nel caso in cui C ha un unico elemento).

Se C ha interno non vuoto anche

$$C - D = \bigcup_{d \in D} (C - d)$$

ha interno non vuoto grazie alla *continuità delle traslazioni,* per ogni insieme D (e quindi anche per $D = C$).

Inoltre, se C contiene punti interni, allora $(\mathrm{cl}\, C) - (\mathrm{cl}\, C)$ contiene un intorno di 0.

Proviamo ora il teorema 102. Ricordiamo che *per definizione, un operatore $A \in \mathcal{L}(X, Y)$ ha dominio uguale ad X.*

Per provare che l'operatore A, suriettivo, trasforma aperti in aperti, è sufficiente mostrare che l'immagine di una palla

$$B_{X,r} = \{x \in X \mid ||x|| < r\}$$

contiene una palla $B_{Y,\sigma}$,

$$B_{Y,\sigma} = \{y \in Y \mid ||y|| < \sigma\}\,.$$

Ciò prova che $A0$ è interno ad $\mathrm{im}\, A$ e, per traslazione, si trova che Ax_0 è interno ad $\mathrm{im}\, A$ per ogni x_0 (nuovamente, si usa la *continuità delle traslazioni*).

Precisamente, proveremo che esiste un intorno W di 0 in Y che è contenuto in $A(B_{X,2})$. Useremo per questo l'inclusione seguente, che vale per ogni $r > 0$ ed $r' > 0$:

$$B_{X,r} - B_{X,r'} \subseteq B_{X,r+r'}\,. \tag{2.43}$$

Consideriamo le palle

$$B_{X,2^n} = \{x \in X \mid ||x|| < 2^n\}\,.$$

Dato che

$$X = \bigcup_n B_{X,2^n}$$

e che A è suriettivo, si trova

$$Y = \bigcup_n \mathrm{cl}\ (A(B_{X,2^n}))$$

e quindi, per il *teorema di Baire* almeno uno degli insiemi $\mathrm{cl}\ (A(B_{X,2^n}))$ ha punti interni. Moltiplicando per numeri positivi si vede che ciascuno degli insiemi

cl $(A(B_{X,r}))$ contiene punti interni, grazie alla *continuità della moltiplicazione per scalari.*

Da (2.43) si ha che

$$\mathrm{cl}\, A(B_{X,r}) - \mathrm{cl}\, A(B_{X,r}) \subseteq \mathrm{cl}\, [A(B_{X,r}) - A(B_{X,r})] \subseteq \mathrm{cl}\, A(B_{X,2r})$$

e quindi *ciascun insieme* $\mathrm{cl}\, A(B_{X,2r})$ *contiene un intorno di* 0. Naturalmente, r è arbitrario: ogni insieme $\mathrm{cl}\, A(B_{X,r})$ contiene un opportuno intorno di 0.

Rimane da provare che $A(B_{X,r})$ stesso contiene un intorno di 0. Sia W un intorno di 0 contenuto in $\mathrm{cl}\, A(B_{X,1})$. Completiamo la dimostrazione mostrando che $W \subseteq A(B_{X,2})$. Per questo basta provare

$$\mathrm{cl}\, A(B_{X,1}) \subseteq A(B_X, 2)\,.$$

Sia $\tilde{y} \in \mathrm{cl}\, A(B_{X,1})$. Mostriamo che $\tilde{y} \in A(B_{X,2})$. Si sa che $\mathrm{cl}\, A(B_{X,1/2})$ contiene un intorno di 0. Dunque esiste $x_1 \in B_{X,1}$ tale che $||\tilde{y} - Ax_1||$ è così piccolo da aversi $\tilde{y} - Ax_1 \in \mathrm{cl}\, A(B_{X,1/2})$.

In modo analogo, $\mathrm{cl}\, AB_{X,1/4}$ contiene un intorno di 0 e quindi esiste $x_2 \in B_{X,1/2}$ per cui $(\tilde{y} - Ax_1) - Ax_2 \in \mathrm{cl}\, A(B_{X,1/4})$. Iterando questo procedimento per ogni n si trova

$x_1 \in B_{X,1}$	tale che	$\tilde{y} - Ax_1 \in \mathrm{cl}\, A(B_{X,1/2})$	cioè	$\lvert\lvert\tilde{y} - Ax_1\rvert\rvert \le 1/2$
$x_2 \in B_{X,1/2}$	tale che	$(\tilde{y} - Ax_1) - Ax_2 \in \mathrm{cl}\, A(B_{X,1/4})$	cioè	$\lvert\lvert(\tilde{y} - Ax_1) - Ax_2\rvert\rvert \le 1/4$
$\vdots$				
$x_n \in B_{X,1/2^n}$	tale che	$\tilde{y} - \sum_{k=1}^n Ax_n \in \mathrm{cl}\, A(B_{X,1/2^n})$	cioè	$\lvert\lvert\tilde{y} - \sum_{k=1}^n Ax_n\rvert\rvert \le 1/2^n\,.$

Sia ora

$$x = \sum_{i=1}^{+\infty} x_i \qquad \text{così che} \qquad ||x|| \le \sum_{i=1}^{+\infty} \frac{1}{2^n} < 2\,.$$

Per questo vettore x vale

$$||\tilde{y} - Ax|| = \lim ||\tilde{y} - Ax_n|| \le \lim \frac{1}{2^n} = 0\,, \quad \text{ossia} \quad \tilde{y} = Ax \quad \text{con} \quad ||x|| < 2\,.$$

Ciò mostra che ogni $\tilde{y} \in \mathrm{cl}\, A(B_{X_1})$ è anche in $A(B_{X,2})$ e conclude la dimostrazione. ∎

2.8.3.1 Le dimostrazioni dei risultati dell'Appendice 2.8.2

Dimostrazione del TEOREMA 115.

Consideriamo in $C(0,1)$ il s.insieme A_n i cui elementi sono funzioni f con questa proprietà: *esiste* $x \in [0, 1 - \frac{1}{n}]$ ed *esiste* $h \in (0, 1-x)$ tale che

$$|f(x+h) - f(x)| \le nh\,.$$

Si prova che:

- L'insieme A_n è chiuso e privo di punti interni.

Accettando queste proprietà che proveremo più avanti, il Teorema di Baire mostra che esiste una funzione continua $f(x)$ che *non appartiene* a $\cup A_n$.

Se $f \notin \cup A_n$ allora *per ogni* x e *per ogni* h vale

$$|f(x+h) - f(x)| > nh$$

e ciò *per ogni* n; ossia il rapporti incrementale è illimitato e quindi la derivata $f'(x)$ non esiste, e ciò per ogni x.

Per completare la dimostrazione, mostriamo che gli insiemi A_n sono chiusi e privi di punti interni.

Proviamo prima di tutto che A_n è chiuso. Sia per questo $f_k \to f$ (uniformemente su $[0,1]$), con $f_k \in A_n$. Dunque, esiste $x_k \in [0, 1-1/n]$ tale che

$$|f_k(x_k+h) - f_k(x_k)| \le nh\,.$$

Passando ad una s.successione, si può assumere $x_k \to x_0 \in [0, 1-1/n]$. Vale:

$$\begin{aligned}
&|f(x_0+h) - f(x_0)| \\
&\le |f(x_0+h) - f_k(x_0+h)| \qquad (2.44)\\
&+|f_k(x_0+h) - f_k(x_k+h)| \qquad (2.45)\\
&+|f_k(x_k+h) - f_k(x_k)| \qquad (2.46)\\
&+|f_k(x_k) - f_k(x_0)| \qquad (2.47)\\
&+|f_k(x_0) - f(x_0)|\,. \qquad (2.48)
\end{aligned}$$

Il termine (2.46) verifica

$$|f_k(x_k+h) - f_k(x_k)| \le nh$$

Essendo f limite uniforme di f_k, per k sufficientemente grande i due addendi (2.44) ed (2.48) sono minori di un prefissato $\epsilon > 0$.

Usiamo ora il Teorema di Ascoli-Arzelà: essendo convergente, la successione (f_k) è equicontinua e: $||(x_0 + h) - (x_k + h)|| \to 0$. Dunque, per k grande, anche gli addendi (2.45) e (2.47) sono minori di ϵ.

Ricapitolando, la funzione f verifica, per ogni $\epsilon > 0$,

$$|f(x_0+h)-f(x_0)| \le nh+4\epsilon \quad \text{e, essendo } \epsilon \text{ arbitrario,} \quad |f(x_0+h)-f(x_0)| \le nh\,.$$

Ciò prova che ciascuno degli insiemi A_n è chiuso. Proviamo ora che ciascuno di essi è privo di punti interni. Fissato n ed $f \in A_n$, proviamo che per ogni $\epsilon > 0$ esiste $\zeta \notin A_n$ che dista da f meno di ϵ.

Si sa che esistono funzioni g continue e lineari a tratti tali che

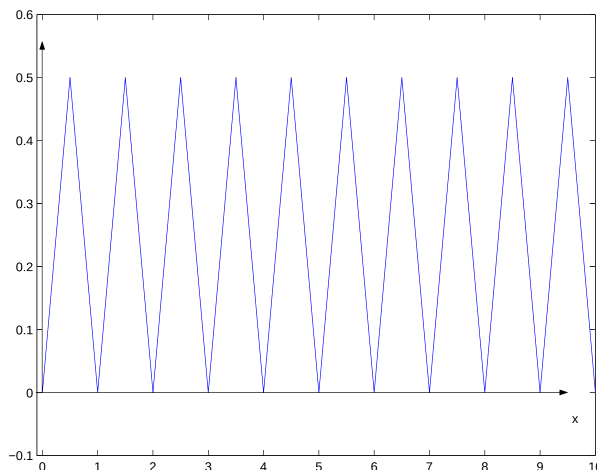
$$||f - g|| < \epsilon/2\,.$$

Basta quindi provare che data una qualunque g continua e lineare a tratti si può costruire $\zeta \notin A_n$, che dista meno di $\epsilon/2$ da g. Sia per questo

$$\phi(x) = \text{distanza di } x \text{ dall'intero più vicino.}$$

Il grafico di $\phi(x)$ è in figura 2.4:

Figura 2.4:



Fissiamo quindi una funzione g lineare a tratti. Essa è lipschitziana e quindi soddisfa $|g(x) - g(x')| < r|x - x'|$ per un r opportuno. Sia ζ la funzione

$$\zeta(x) = g(x) + \epsilon\phi(mx)\,.$$

Chiaramente, per ogni m, $||g - \zeta|| \leq \epsilon/2$. Vogliamo mostrare che, per un'opportuna scelta di m, $\zeta \notin A_n$. Fissiamo x e quindi x' tali che $|x - x'| < 1/2m$. Si ha

$$|mx - mx'| < 1/2\,, \qquad |\phi(mx) - \phi(mx')| = m|x - x'|$$

e quindi

$$\begin{aligned}|\zeta(x) - \zeta(x')| &\geq \Big| \epsilon|\phi(mx) - \phi(mx')| - |g(x) - g(x')| \Big| \\ &= \Big| \epsilon m|x - x'| - |g(x) - g(x')| \Big| \geq (\epsilon m - r)|x - x'|\end{aligned}$$

se $m > r/\epsilon$. Se ora m verifica anche $m > (n + r)/\epsilon$, allora $\zeta \notin A_n$. Ciò completa la dimostrazione. ∎

<u>Dimostrazione del TEOREMA 116.</u>

Indichiamo con $C_P(-\pi, \pi)$ il s.spazio di $C(-\pi, \pi)$ i cui elementi sono funzioni continue che verificano

$$f(-\pi) = f(\pi)\,.$$

Si vede facilmente che questo è un s.spazio chiuso di $C(-\pi, \pi)$, e quindi esso stesso uno spazio di Banach.

Studiamo le somme parziali della serie di Fourier. Indichiamo per questo con F_N l'operatore definito su $C_P(-\pi, \pi)$ da

$$(F_N f)(x) = \sum_{n=-N}^{N} f_n e^{inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \sum_{n=-N}^{N} e^{in(x-s)}\, \mathrm{d}s\,.$$

Poniamo

$$D_N(x) = \sum_{n=-N}^{N} e^{inx}\,.$$

Si proverà alla fine di questo paragrafo che

$$D_N(x) = \begin{cases} 2N + 1 & \text{se } x = 0 \\ \frac{\sin(N+1/2)x}{\sin x/2} & \text{se } x \neq 0\,. \end{cases} \tag{2.49}$$

La funzione $D_N(x)$ si chiama NUCLEO DI DIRICHLET.

Dunque,

$$(F_N f)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) D_N(x - s)\, \mathrm{d}s = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \frac{\sin[(N + 1/2)(x - s)]}{\sin[(x - s)/2]}\, \mathrm{d}s\,.$$

Si osservi che la seconda uguaglianza vale perché $D_N(x)$ è continua e quindi il valore di $D_n(0)$ implicitamente compare nel secondo integrale. Inoltre, la continutà di $D_N(x)$ mostra che ciascun operatore F_N da $C_P(-\pi, \pi)$ in sé è continuo.

Facciamo vedere che esiste una funzione $\tilde{f} \in C_P(-\pi, \pi)$ tale che

$$\sup_N |(F_N \tilde{f})(x_0)| = +\infty\,.$$

Ciò vuol dire che la serie di Fourier di questa funzione $\tilde{f}$ *non* converge in x_0.

Per provare l'esistenza di $\tilde{f}$ si procede per assurdo. Supponiamo che tale funzione $\tilde{f}$ non esista. Allora, *per ogni* $f \in C_P(-\pi, \pi)$ si ha:

$$\sup_N |(F_N f)(x_0)| < +\infty\,.$$

In tal caso *per ogni* $f \in C_P(-\pi, \pi)$ *esiste* M_f per cui

$$|(F_N f)(x_0)| < M_f$$

per ogni N. E quindi, per il *teorema di Banach–Steinhaus,* esiste $M = M(x_0)$, *indipendente da* N*,* tale che

$$|(F_N f)(x_0)| < M(x_0)||f||_{C_P(-\pi,\pi)}\,.$$

Indicando con $F_{x_0,N}$ il funzionale che ad $f \in C_P(-\pi, \pi)$ associa $(F_N f)(x_0)$, la disuguaglianza precedente si scrive

$$||F_{x_0,N}|| < M(x_0)\,; \tag{2.50}$$

ossia, la famiglia dei funzionali lineari e continui $F_{x_0,N}$ è limitata. Calcoliamo esplicitamente la norma del funzionale $F_{x_0,N}$ e mostriamo che ciò non vale.

Per semplicità limitiamoci a fare il calcolo con $x_0 = 0$. In questo caso

$$F_{0,N} f = \frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(t) f(t)\, \mathrm{d}t\,.$$

e

$$|F_{0,N} f| \le \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(t) f(t)|\, \mathrm{d}t \le \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(t)|\, \mathrm{d}t \right) ||f||_{C_P(-\pi,\pi)}$$

così che

$$||F_{0,N}|| \le \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(t)|\, \mathrm{d}t < +\infty \quad \text{perché } D_N(x) \text{ è continuo}\,. \tag{2.51}$$

In realtà vedremo che vale l'uguaglianza. Accettando ciò,

$$
\begin{aligned}
&||F_{0,N}|| = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_N(t)|\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\frac{\sin[(N+1/2)t]}{\sin t/2}\,\mathrm{d}t\right| \\
&> \frac{1}{\pi}\int_{-\pi}^{\pi} \left|\frac{\sin[(N+1/2)t]}{t}\right|\,\mathrm{d}t = \frac{2}{\pi}\int_0^{(N+1/2)\pi} \frac{|\sin t|}{t}\,\mathrm{d}t \\
&\geq \frac{2}{\pi}\sum_{k=0}^{N-1}\int_{k\pi}^{(k+1)\pi} \frac{|\sin t|}{t}\,\mathrm{d}t \geq \sum_{k=0}^{N-1} \frac{2}{(k+1)\pi^2}\int_{k\pi}^{(k+1)\pi} |\sin t|\,\mathrm{d}t \\
&= \frac{4}{\pi^2}\sum_{k=0}^{2n} \frac{1}{(k+1)} \longrightarrow +\infty\,.
\end{aligned}
$$

Ciò contrasta con la (2.50) e mostra che la funzione $\tilde{f}$ *esiste.*

Accenniamo ora alla dimostrazione del fatto che l'uguaglianza vale nella formula (2.51). Per mostrare ciò è sufficiente trovare una successione di funzioni (f_k) di norma al più uguale ad 1 e tale che

$$
\lim_k |F_{0,N} f_k| = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_N(t)|\,\mathrm{d}t \tag{2.52}
$$

Introduciamo per questo la funzione

$$
y(x) = \operatorname{sign} D_N(x)
$$

e una successione (f_k) di funzioni continue convergente puntualmente ad y e inoltre limitata da 1. L'uguaglianza (2.52) vale per questa successione di funzioni.

Ciò completa la dimostrazione, a parte il fatto che dobbiamo ancora verificare l'uguaglianza (2.49), cosa che facciamo ora.

L'uguaglianza è ovvia se x è multiplo di 2π. Altrimenti, usando la formula

per la somma di una progressione geometrica, si ha:

$$
\begin{aligned}
D_N(x) &= \sum_{k=-N}^{N} e^{ikx} = \sum_{k=-N}^{-1} e^{ikx} + \sum_{k=0}^{N} e^{ikx} = \sum_{k=1}^{N} e^{-ikx} + \sum_{k=0}^{N} e^{ikx} \\
&= \left[\frac{1-e^{-i(N+1)x}}{1-e^{-ix}} - 1 \right] + \frac{1-e^{i(N+1)x}}{1-e^{ix}} \\
&= \frac{e^{-ix} - e^{-iNx}e^{-ix}}{1-e^{-ix}} + \frac{1-e^{i(N+1)x}}{1-e^{ix}} \\
&= \frac{\left[e^{-ix} - e^{-iNx}e^{-ix}\right](1-e^{ix}) + \left[1-e^{i(N+1)x}\right](1-e^{-ix})}{(1-\cos x)^2 + \sin^2 x} \\
&= \frac{e^{-iNx}(1-e^{-ix}) + e^{iNx}(1-e^{ix})}{2-2\cos x} = \frac{2\Re e\left[e^{-iNx}(1-e^{-ix})\right]}{2-2\cos x} \\
&= \frac{\cos Nx - \cos Nx \cos x + \sin x \sin Nx}{4\sin^2(x/2)} \\
&= \frac{\cos Nx - \cos(N+1)x}{2\sin^2(x/2)} = \frac{2\sin(x/2)\sin\frac{1}{2}(2N+1)x}{2\sin^2(x/2)} \\
&= \frac{\sin\left(N+\frac{1}{2}\right)x}{\sin\frac{x}{2}} .
\end{aligned}
$$

Nel penultimo passaggio si sono usate le formule di prostaferesi. ∎

2.9 Lo spazio duale

Abbiamo già visto la relazione tra i funzionali lineari e continui su X e la nozione geometrica di iperpiano. Ciò suggerisce di studiare più a fondo i funzionali lineari continui, sia singolarmente che nel loro insieme, studiando le proprietà dello spazio di Banach X^*.

E' ovvio che il funzionale 0, quello che ad ogni elemento di X associa l'elemento nullo del campo scalare, è in X^*. Non è affatto ovvio che esistano altri elementi di X^*. Infatti:

Teorema 117 *Esistono spazi lineari X, dotati di una metrica rispetto alla quale le operazioni di somma e moltiplicazioni per scalari sono continue e su cui nessun funzionale lineare diverso da 0 è continuo.*

Ossia, in tal caso, nello spazio X non si trovano iperpiani.

La dimostrazione è posposta.

E' quindi estremamente importante sapere che se X è uno spazio di Banach allora X^* ha "molti" elementi. Ciò è conseguenza del teorema seguente:

Teorema 118 (di HAHN-BANACH) *Sia X uno s.l.n. (anche non completo) e sia Y un suo s.spazio. Sia L_0 un funzionale* lineare continuo *su Y. Esiste un'estensione di L ad X tale che*

$$||L||_{X^*} = \sup\{|Ly| \mid y \in Y\,,\ ||y||_X = 1\}\,.$$

Ossia, ogni funzionale lineare e continuo su Y può estendersi ad X senza alterarne la norma.

Si noti che nell'enunciato precedente si può assumere che Y sia chiuso, perché l'operatore L_0 si può estendere per continuità alla chiusura del suo dominio.

Per certe applicazioni (allo studio degli insiemi convessi, si veda il paragrafo 2.9.1) è necessario provare una versione un po' più generale del Teorema 118; ossia è necessario provare i due teoremi seguente:

Teorema 119 *Sia X uno spazio lineare su $\mathbb{R}$. Esista una funzione $p\colon X \to \mathbb{R}$*

1. POSITIVAMENTE OMOGENEA, *ossia tale che $p(tx) = tp(x)$ per ogni x e per ogni $t \geq 0$;*
2. SUBADDITIVA, *ossia tale che $p(x+y) \leq p(x) + p(y)$ per ogni x, y in X.*

Sia Y un s.spazio di X e sia L_0 un funzionale lineare definito su Y, tale che

$$L_0 x \leq p(x) \qquad \forall x \in Y\,.$$

Esiste un'estensione L di L_0 ad X che verifica

$$Lx \leq p(x) \qquad \forall x \in X\,.$$

Teorema 120 *Sia X uno spazio lineare su $\mathbb{C}$ e sia p da X in $\mathbb{R}$ una funzione tale che:*

1. *$p(tx) = |t| \cdot p(x)$ per ogni x in X e per ogni $t \in \mathbb{C}$;*
2. *$p(x+y) \leq p(x) + p(y)$.*

Sia Y un s.spazio di X e sia L_0 un funzionale lineare definito su Y, tale che

$$|L_0 x| \leq p(x) \qquad \forall x \in Y\,. \tag{2.53}$$

Esiste un'estensione L di L_0 ad X che verifica

$$|Lx| \leq p(x) \qquad \forall x \in X\,.$$

Posponiamo le dimostrazioni, notando che la dimostrazione del Teorema 120 si ridurrà, con un opportuno artificio, a quella del Teorema 119.

Dimostrazione del TEOREMA 118. La dimostrazione discende immediatamente dai Teoremi 119 e 120. Sia

$$M = \sup\{|L_0 x| \mid x \in Y\,,\ ||x|| = 1\}\,, \qquad p(x) = M||x||\,.$$

Nel caso reale, dal Teorema 119 si vede l'esistenza di L, funzionale lineare su X, che estende L_0 e tale che

$$Lx \leq M||x||$$

e quindi anche

$$-Lx = L(-x) \leq M||-x|| = M||x||$$

ossia

$$|Lx| \leq M||x||\,. \tag{2.54}$$

Nel caso complesso la disuguaglianza (2.54) figura direttamente nell'enunciato del Teorema 120. Dunque il Teorema 118 vale. ∎

Mostriamo ora alcune conseguenze importanti del Teorema di Hahn-Banach. La prima è che X^*, duale di X, ha "molti" elementi. Più precisamente,

Teorema 121 *Per ogni $x_0 \neq 0$ in X ed ogni $m \in \Phi$, esiste $L \in X^*$ tale che*

$$Lx_0 = m\,, \qquad ||L|| = \frac{m}{||x_0||}\,. \tag{2.55}$$

Dim. Si sceglie come spazio Y la retta per 0 ed x_0,

$$Y = \{\lambda x_0 \mid \lambda \in \Phi\}\,.$$

Il s.spazio ha dimensione 1 e su esso è facile definire

$$L_0(\lambda x_0) = \lambda m\,.$$

Il funzionale L_0 verifica (2.55) sui soli elementi di Y. Si usa quindi il Teorema di Hahn-Banach per estendere L ad X senza alterarne la norma. ∎

In particolare si può segliere $m = 1$ oppure $m = ||x_0||$ oppure $m = ||x_0||^2$. In particolare, con quest'ultima scelta si trova

$$||L|| = ||x_0||\,.$$

Invece, scegliendo $m = ||x_0||$, si trova

Corollario 122 *Sia $x_0 \in X$. Vale:*

$$||x_0|| = \max_{||L||_{X^*}=1} |Lx_0|\,.$$

<u>Dim.</u> Si fissi $x_0 \neq 0$ in X. Per ogni $L \in X^*$ di norma 1 si ha:

$$|Lx_0| \leq ||x_0||\,, \qquad \text{ossia} \qquad \sup_{||L||_{X^*}=1} ||Lx_0|| \leq ||x_0||\,.$$

Il funzionale definito in (2.55) con $m = ||x_0||$ ha norma 1 e per esso

$$|Lx_0| = ||x_0||\,. \blacksquare$$

Osservazione 123 L'asserto del corollario precedente somiglia alla definizione della norma di un funzionale,

$$||L|| = \sup_{||x||=1} |Lx| \tag{2.56}$$

Ma abbiamo provato che *in generale, l'estremo superiore in 2.56* non *è un massimo,* si veda il teorema 80.

Ciò mostra una prima differenza importante tra uno spazio ed il suo duale: lo spazio duale "si comporta meglio" dello spazio X stesso. Avremo altre occasioni di notare questo fatto, in particolare quando proveremo il Teorema di Banach-Alaoglu, Teorema 158. ∎

Il Teorema 121 in particolare afferma che se un elemento x_0 è diverso da 0 allora esiste un funzionale L di X^* che "lo vede non nullo"; e, traslando, se $x_1 \neq x_0$, esiste un $L \in X^*$ tale che

$$Lx_1 \neq Lx_0\,.$$

Dunque, X^* ha così tanti elementi da distinguere quelli di X. Geometricamente, abbiamo provato l'esistenza di un iperpiano che non contiene ambedue gli elementi x_0 ed x_1. Questa osservazione può essere estesa fino a "separare" mediante iperpiani due insiemi convessi tra loro disgiunti. Prima di studiare gli insiemi convessi conviene però introdurre una notazione comoda per indicare l'azione degli elementi di X^* su X. Invece di scrivere Lx, Ax ecc, usa scrivere[16]

$$\langle\langle f, x\rangle\rangle$$

[16]In realtà la notazione comunemente usata è $\langle\cdot,\cdot\rangle$. Noi usiamo la notazione $\langle\langle\cdot,\cdot\rangle\rangle$ perché la notazione $\langle\cdot,\cdot\rangle$ si usa anche per indicare i "prodotti interni" nel contesto degli spazi di Hilbert. Dato che vedremo una relazione tra funzionali lineari e prodotti interni, è opportuno essere precisi nel distinguere gli uni dagli altri.

per indicare il valore che il funzionale $f \in X^*$ assume sull'elemento $x \in X$ (si noti: *l'elemento di X^* è scritto prima di quello di X.* In altri testi si trova scritto dopo). Inoltre, per distinguere immediatamente gli elementi di X da quelli di X^*, usa indicare questi ultimi con lettere greche, o con simboli del tipo x^*, y^* ecc. se i simboli x, y,... si riservano agli elementi di X.

2.9.1 Applicazioni: Insiemi convessi

Per semplicità *supponiamo che il campo scalare sia* $\mathbb{R}$.

Ricordiamo che un insieme A non vuoto si dice CONVESSO quando ogni segmento di estremi in A è contenuto in A; ossia quando[17]

$$x\,,\ y \in A \Longrightarrow tx + (1-t)y \in A \qquad \forall t \in [0,1]\,.$$

Si verifica facilmente che ogni s.spazio V è convesso, e quindi anche $x_0 + V$ è convesso; gli IPERPIANI, ossia gli insiemi

$$\{x\,:\langle\langle x^*, x\rangle\rangle \geq 0\} \qquad x^* \in X^*\,,$$

sono convessi.

Siano ora A e B due insiemi e sia $x^* \in X^*$. Si dice che l'iperpiano[18]

$$\{x \mid \langle\langle x^*, x\rangle\rangle = \alpha\}$$

SEPARA i due insiemi A e B se

$$A \subseteq \{x \mid \langle\langle x^*, x\rangle\rangle \leq \alpha\}\,, \qquad B \subseteq \{x \mid \langle\langle x^*, x\rangle\rangle \geq \alpha\}\,.$$

Si dice che la separazione è *stretta* se esiste $\epsilon > 0$ tale che

$$A \subseteq \{x \mid \langle\langle x^*, x\rangle\rangle \leq \alpha - \epsilon\}\,, \qquad B \subseteq \{x \mid \langle\langle x^*, x\rangle\rangle \geq \alpha\}\,.$$

E' ovvio che in generale due insiemi disgiunti non possono essere separati da iperpiani, si veda la figura 2.5 a sinistra. Per avere buoni risultati di separazione dovremo lavorare con insiemi convessi. La figura 2.5 a destra mostra che insiemi convessi e disgiunti non possono, in generale, separarsi strettamente.

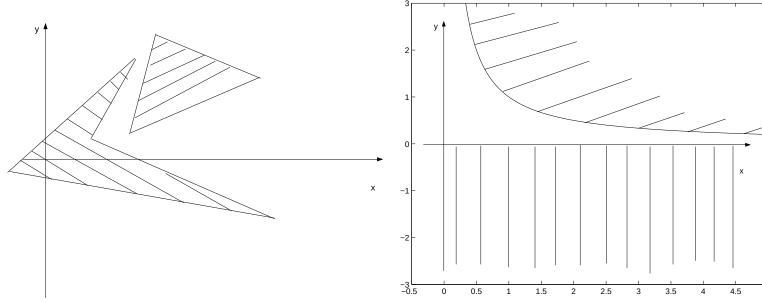
Valgono però i due teoremi seguenti, la cui dimostrazione viene posposta:

---

[17] non si richiede $x \neq y$ e quindi se A contiene un solo punto esso è convesso. Conviene considerare convesso anche l'insieme vuoto, per semplificare certi enunciati. Per esempio, *l'intersezione di insiemi convessi è un convesso* vale se $\emptyset$ si definisce come convesso.

[18] si noti che le disuguaglianze $\leq$ per A e $\geq$ per B sono puramente convenzioanli e possono scambiarsi cambiando il segno di x^*.

Figura 2.5:



Teorema 124 *Sia $A \subseteq X$ un convesso aperto. Se $x_0 \notin A$ allora esiste un iperpiano che separa x_0 da A; ossia esiste $x^* \in X^*$ tale che*

$$\langle\langle x^*, x\rangle\rangle \leq \langle\langle x^*, x_0\rangle\rangle \qquad \forall x \in A\,.$$

e

Teorema 125 *Siano A, B convessi e disgiunti. Si ha:*

- *Se A è* aperto, *esiste un iperpiano che* separa *A e B;*
- *se A è* compatto *e B è* chiuso *allora esiste un iperpiano che* separa strettamente *A e B.*

Di conseguenza:

Teorema 126 *Sia H un s.spazio chiuso dello spazio di Banach X. Vale $H \neq X$ se e solo se esiste $x^* \in X^*$, $x^* \neq 0$, tale che*

$$H \subseteq \ker x^*\,.$$

<u>Dim</u>. Se $H \neq X$, esiste $x_0 \notin H$ e quindi per la seconda affermazione del Teorema 125, x_0 si separa strettamente da H: esistono $\epsilon > 0$ ed x^* tale che

$$\langle\langle x^*, x_0\rangle\rangle \leq \alpha - \epsilon\,, \qquad \langle\langle x^*, h\rangle\rangle \geq \alpha\,,\ \forall h \in H\,.$$

Essendo H un s.spazio, se $h \in H$ allora si ha anche $-h \in H$ e quindi

$$\alpha \leq \langle\langle x^*, -h\rangle\rangle = -\langle\langle x^*, h\rangle\rangle\,.$$

Dunque $\alpha = 0$ e

$$\langle\langle x^*, h\rangle\rangle = 0 \qquad \forall h \in H\,,$$

come volevamo.

Il viceversa è ovvio. ∎

Inoltre,

Teorema 127 *Se X^* è separabile, anche X lo è.*

Dim. Sia $\{x_n^*\}$ un s.insieme denso in X^*. Si scelgano elementi $x_n \in X$ di norma 1 e tali che

$$\left|\langle\langle x_n^*, x_n \rangle\rangle\right| \geq \frac{1}{2}\lambda_n\,, \qquad \lambda_n = ||x_n^*||\,.$$

Combinando linearmente gli x_n, con coefficienti razionali, si trova un s.insieme A numerabile di X. Procedendo per assurdo, mostriamo che A è denso in X, che pertanto è separabile.

Se A non fosse denso in X, la sua chiusura sarebbe un s.spazio H di X *diverso* da X. Dunque, per il Teorema 126, si potrebbe trovare $\phi^* \in X^*$, $\phi^* \neq 0$, tale che

$$A \subseteq H \subseteq \ker \phi^*\,.$$

Si sa che $\phi^* = \lim x_{n_k}^*$ per una opportuna successione $(x_{n_k}^*)$ così che

$$\begin{aligned} 0 = \lim ||\phi^* - x_{n_k}^*||_{X^*} &\geq \limsup \left|\langle\langle \phi^* - x_{n_k}^*, x_{n_k} \rangle\rangle\right| \\ = \limsup \left|\langle\langle x_{n_k}^*, x_{n_k} \rangle\rangle\right| &\geq \limsup \frac{1}{2}\lambda_n \end{aligned}$$

e quindi $\lim \lambda_n = 0$. D'altra parte, mentre

$$0 \neq ||\phi^*||_{X^*} = \lim ||x_{n_k}^*||_{X^*} = \lim \lambda_{n_k}\,.$$

La contraddizione trovata completa la dimostrazione. ∎

Osservazione 128 Invece, *il duale di uno spazio separabile può non essere separabile.* Mostreremo infatti che l^∞ è isometricamente isomorfo al duale di l^1. Si sa già che l^∞ non è separabile mentre è facile verificare che l^1 lo è. ∎

Un'ulteriore conseguenza importante del Teorema 125 è la seguente: ogni insieme A, anche non convesso, è contenuto nell'intersezione dei semispazi che lo contengono. Se A è convesso vale:

Teorema 129 *Un insieme A non vuoto è convesso e chiuso se e solo se è intersezione dei semispazi chiusi che lo contengono.*

Dim. L'intersezione di una famiglia di chiusi è un chiuso e l'intersezione di una famiglia di convessi è un convesso. Dunque, se l'insieme non vuoto A è intersezione dei semispazi chiusi che lo contengono, A è convesso e chiuso.

Viceversa, sia B l'intersezione dei semispazi chiusi che contengono l'insieme A. In generale, $B \supseteq A$. Dobbiamo provare che se l'insieme non vuoto A è convesso e chiuso allora vale l'uguaglianza. Per questo basta notare che se $x_0 \notin A$ allora, per il Teorema 125, esiste un iperpiano che separa strettamente A da x_0. E quindi x_0 non appartiene nemmeno a B. ∎

Si chiamano IPERPIANI DI SUPPORTO all'insieme convesso e chiuso A quelli che godono della seguente proprietà: esiste $x_0 \in \operatorname{cl} A$ tale che

$$A \subseteq \{x \mid \langle\langle x^*, x\rangle\rangle \le \langle\langle x^*, x_0\rangle\rangle\}\,.$$

L'iperpiano di supporto è

$$\{x : \quad \langle\langle x^*, x\rangle\rangle - \langle\langle x^*, x_0\rangle\rangle\,.$$

Più precisamente, in questo caso si dice che l'iperpiano è di supporto nel punto x_0.

Naturalmente, se esiste un iperpiano di supporto ad A in x_0 allora x_0 non è interno ad A. Più precisamente:

Teorema 130 *Sia A un convesso con interno non vuoto. Per ogni $x_0 \in \partial A$, esiste almeno un iperpiano di supporto ad A in x_0.*

Dim. Usiamo la proprietà seguente degli insiemi convessi, che non proviamo: *se A è convesso ed il suo interno è non vuoto allora l'interno di A è convesso e denso in A.*

Dunque $\operatorname{int} A$ è un aperto convesso, ovviamente separato da $x_0 \in \partial A$: per il Teorema 124, esiste x^* tale che

$$\langle\langle x^*, x\rangle\rangle \le \langle\langle x^*, x_0\rangle\rangle \qquad \forall x \in \operatorname{int} A$$

e la disuguaglianza si estende per continuità ad A.

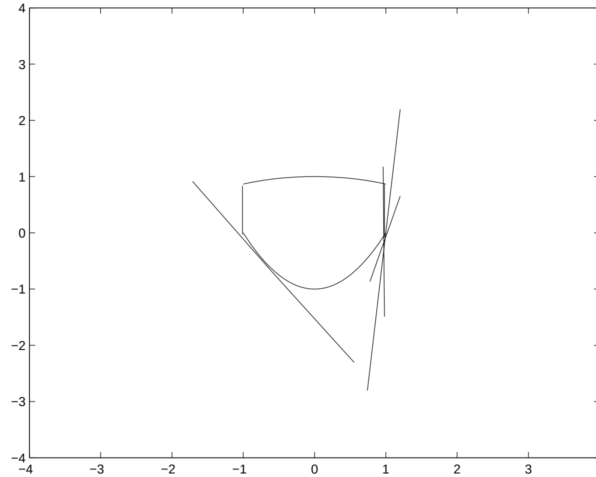
L'iperpiano di supporto è

$$\{x \mid \langle\langle x^*, x\rangle\rangle = \langle\langle x^*, x_0\rangle\rangle\}\,. \;∎$$

Osservazione 131

- Il teorema precedente può anche estendersi al caso in cui l'interno di A è vuoto. Essendo A convesso, si prova l'esistenza di un più piccolo s.spazio (traslato) contenente A. Il teorema si può enunciare in tale s.spazio. Non insistiamo su ciò.

Figura 2.6:



- La figura 2.6 mostra che uno stesso iperpiano può essere di supporto in infiniti punti, e che in un punto si possono avere infiniti iperpiani di supporto. ∎

Abbiamo dunque due modi di descrivere un insieme convesso e chiuso, tra loro equivalenti: il primo consiste nell'elencare i punti dell'insieme ed il secondo consiste nell'elencare gli iperpiani di supporto all'insieme. Questo secondo modo assume un aspetto particolare nel caso in cui l'insieme A è l'epigrafo di una funzione, caso che ora andiamo a studiare.

2.9.2 Applicazioni: Funzioni convesse

Ancora, *supponiamo di lavorare in spazi lineari su* $\mathbb{R}$ e studiamo certe proprietà delle funzioni convesse definite su uno spazio di *Banach* X. Per semplicità supporremo che esse siano definite su insiemi, ovviamente convessi, *dotati di punti interni.*

Ricordiamo che si chiama EPIGRAFO di una funzione reale f l'insieme

$$\text{Epi}(f) = \{(x,t) \mid x \in \text{dom}\, f\,, \quad t \geq f(x)\} \subseteq X \times \mathbb{R}\,.$$

E' facile provare:

Teorema 132 *La funzione f è convessa se e solo se il suo epigrafo è un insieme convesso.*

Un funzionale di supporto in (x_0, t_0) alla chiusura dell'epigrafo di f è una coppia $(x^*, \phi) \in X^* \times \mathbb{R}$ tale che, per ogni punto (x, t) dell'epigrafo, valga

$$\langle\langle x^*, x\rangle\rangle + \phi \cdot t \le \langle\langle x^*, x_0\rangle\rangle + \phi \cdot t_0$$

ossia

$$\langle\langle x^*, (x - x_0)\rangle\rangle \le (-\phi) \cdot (t - t_0)\,. \qquad (2.57)$$

Notiamo che se $\phi = 0$ si ha

$$\langle\langle x^*, x\rangle\rangle \le \langle\langle x^*, x_0\rangle\rangle \qquad \forall x \in \mathrm{dom}\, f\,.$$

Dunque:

Lemma 133 *Se*

$$(x^*, 0) \in X^* \times \mathbb{R}$$

è funzionale di supporto in (x_0, t_0) alla chiusura dell'epigrafo di f allora $x_0 \in \partial \mathrm{dom}\, f$. Quindi, se x_0 è interno a $\mathrm{dom}\, f$ un iperpiano di supporto in (x_0, t) deve avere $\phi \ne 0$; geometricamente, non è un iperpiano verticale. In tal caso, $-\phi > 0$.

Supponendo di lavorare in *punti interni* al dominio, avremo $\phi \ne 0$ e si ha $-\phi > 0$. Infatti se $x = x_0$ e $t > t_0$ allora deve essere

$$0 \le (-\phi)(t - t_0) \qquad \forall t \ge t_0 \qquad \text{ossia } -\phi > 0.$$

Dunque non è restrittivo assumere $-\phi = 1$ perché dividendo i due membri per $-\phi$ (che è positivo) e rinominando $x^*/(-\Phi)$, si vede che la (2.57) può scriversi

$$\langle\langle x^*, x - x_0\rangle\rangle \le t - t_0\,.$$

Supponiamo ora che l'epigrafo di f sia chiuso. In questo caso $(t_0 f(x_0)) \in \mathrm{Epi}(f)$. Scegliendo $t \ge f(x)$ si ha:

Teorema 134 *Se l'epigrafo di f è chiuso allora il funzionale $(x^*, 1) \in X^* \times \mathbb{R}$ è di supporto all'epigrafo di f in $(x_0, f(x_0))$ se e solo se*

$$\langle\langle x^*, x - x_0\rangle\rangle \le f(x) - f(x_0) \qquad \forall x \in \mathrm{dom} f\,. \qquad (2.58)$$

La (2.58) si scrive anche

$$f(x) \ge \langle\langle x^*, x - x_0\rangle\rangle + f(x_0)\,.$$

La (2.58) e la pratica con i grafici delle funzioni da $\mathbb{R}$ in $\mathbb{R}$ suggerisce:

Definizione 135 Un funzionale x^* che verifica (2.58) si chiama **un** sottodifferenziale di f in x_0.

L'**insieme** dei sottodifferenziali di f in x_0 si chiama **il** SOTTODIFFERENZIALE di f in x_0.

Il sottodifferenziale di f in x_0 si indica col simbolo $\partial f(x_0)$. ∎

Osserviamo che il sottodifferenziale può essere l'insieme vuoto ma, come si è già notato, in tal caso il punto x_0 deve appartenere alla frontiera del dominio di f:

Esempio 136 Sia $X = \mathbb{R}$ e sia $\mathrm{dom} f = [0, +\infty)$, $f(x) = -\sqrt{x}$. Si vede facilmente che $\partial f(0) = \emptyset$. ∎

Invece, *il sottodifferenziale è non vuoto nei punti interni al dominio della funzione convessa f* perché l'epigrafoe è convesso, ammette quindi funzionale di supporto che $-\Phi = 1$ per il Lemma 133.

E' facile immaginare che il sottodifferenziale sia uno strumento utile per il calcolo dei minimi. Infatti,

Teorema 137 *Sia f convessa e sia non vuoto il suo sottodifferenziale in x_0. Il punto x_0 è punto di minimo se e solo se $0 \in \partial f(x_0)$.*

Dim. Immediate, leggendo la (2.58) con $x^* = 0$. ∎

Infine ,notiamo che La proprietà di avere epigrafo chiuso è ovviamente importante nei problemi di ottimizzazione e merita una definizione specifica:

Definizione 138 Una funzione (anche non convessa) con epigrafo chiuso si chiama SEMICONTINUA INFERIORMENTE. Una funzione f si dice *semicontinua superiormente* se $-f$ è semicontinua inferiormente. ∎

2.9.3 Dimostrazioni posposte

Dimostrazione del TEOREMA 117.

Indichiamo con X l'insieme delle (classi di equivalenza di) funzioni $f(x)$ definite su $(0, 1)$ e tali che

$$\int_0^1 \sqrt{|f(x)|}\, \mathrm{d}x < +\infty\,.$$

La disuguaglianza

$$\sqrt{a+b} \le \sqrt{a} + \sqrt{b} \qquad \forall a \ge 0\,,\ b \ge 0 \tag{2.59}$$

permette di provare che X è uno spazio vettoriale e che

$$d(f,g)=\int_0^1\sqrt{|f(x)-g(x)|}\,\mathrm{d}x$$

è una distanza invariante per traslazioni su X (per provare la disuguaglianza triangolare si usa la (2.59)).

Si può provare che lo spazio X è completo.

Lo spazio X ha una proprietà curiosa: *l'unico convesso (non vuoto) di X che è anche aperto è X stesso.* Accettando questa proprietà, che proveremo in seguito, si vede che l'unico funzionale L lineare e continuo su X è quello nullo. Infatti, per ogni $\epsilon>0$, l'insieme $L^{-1}(-\epsilon,\epsilon)$ non è vuoto, perché contiene 0; è convesso perchè L è lineare e, se L è anche continuo, è aperto; e quindi è tutto lo spazio X. Dunque,

$$LX\subseteq\bigcap_{\epsilon>0}(-\epsilon,\epsilon)=\{0\}\,;$$

ossia, L è il funzionale nullo.

Per completare la dimostrazione, consideriamo un aperto V di X, non vuoto e convesso e proviamo che $V=X$. A meno di traslazioni, si può supporre $0\in V$. Dunque può trovarsi $r>0$ tale che

$$B(0,r)=\{x\mid d(x,0)<r\}\subseteq V\,.$$

Fissiamo una qualsiasi $f\in X$ e sia n tale che

$$\frac{1}{\sqrt{n}}d(f,0)<r\quad\text{e quindi }(1/\sqrt{n})f\in B\,.$$

Consideriamo ora la funzione integrale

$$F(x)=\int_0^x\sqrt{|f(s)|}\,\mathrm{d}s\,.$$

Vale $F(0)=0$ ed $F(1)=d(f,0)$. La continuità di F implica l'esistenza di un primo punto x_1 tale che

$$F(x_1)=\int_0^{x_1}\sqrt{|f(s)|}\,\mathrm{d}s=\frac{1}{n}d(f,0)=\frac{1}{n}\int_0^1\sqrt{|f(s)|}\,\mathrm{d}s\,;$$

di un primo punto x_2 tale che

$$\int_{x_1}^{x_2}\sqrt{|f(s)|}\,\mathrm{d}s=\frac{1}{n}d(f,0)\,.$$

Analogamente, si trovano esattamente n punti x_i tali che

$$\int_{x_{i-1}}^{x_i} \sqrt{|f(s)|}\,\mathrm{d}s = \frac{1}{n}d(f,0) \qquad (\text{ove } x_0 = 0).$$

Per $1 \leq i \leq n$ definiamo le funzioni

$$g_i(s) = \begin{cases} nf(s) & \text{se } x_{i-1} < s < x_i \\ 0 & \text{altrimenti.} \end{cases}$$

In questo modo,

$$d(g_i, 0) = \frac{\sqrt{n}}{n}d(f,0) = \frac{1}{\sqrt{n}}d(f,0) < r\,,$$

ossia $g_i \in B(0,r) \subseteq V$. Inoltre,

$$nf = \sum_{i=1}^{n} g_i \qquad \text{ossia} \qquad f = \frac{1}{n}\sum_{i=1}^{n} g_i\,.$$

L'ultima uguaglianza mostra che anche f appartiene al *convesso* V.

L'arbitrarietà di f implica che $V = X$. ∎

Dimostrazione del TEOREMA 119.

La dimostrazione del Teorema 119 fa uso del LEMMA DI ZORN che ora enunciamo. Sia $\mathcal{F}$ un insieme *parzialmente* ordinato. Chiamiamo CATENA ogni s.insieme di $\mathcal{F}$ che è *totalmente* ordinato.

Sia M un qualsiasi s.insieme di $\mathcal{F}$ e sia $z \in \mathcal{F}$. Diciamo che z è un MAGGIORANTE di M se per ogni $m \in M$ vale $m \leq z$; diciamo che z_0 è un ELEMENTO MASSIMALE di M se $z_0 \in M$ e se, inoltre, le condizioni $m \in M$ ed $m \geq z_0$ implicano $m = z_0$.

Un elemento massimale di M confrontabile *con tutti* gli elementi di M si chiama un MASSIMO.

Si noti che un elemento massimale z_0 può non essere un massimo perché M può contenere elementi non confrontabili con z_0. E quindi un insieme M *parzialmente* ordinato può avere più elementi massimali, ma ha *al più* un massimo.

Invece, una catena ha al più un solo elemento massimale che è anche il suo massimo.

Il Lemma di Zorn si enuncia come segue:

Lemma 139 (di ZORN) *Sia $\mathcal{F}$ parzialmente ordinato. Se ogni catena di $\mathcal{F}$ ammette maggioranti, allora esiste in $\mathcal{F}$ un elemento massimale.*

Osservazione 140 Si noti che l'enunciato del Lemma di Zorn non richiede che ogni catena debba avere massimo; ossia, nell'applicare il Lemma di Zorn, basta mostrare l'esistenza in $\mathcal{F}$ di maggioranti delle singole catene. ∎

Possiamo ora provare il Teorema 119. In questo caso $\Phi = \mathbb{R}$.

Notiamo che l'enunciato del teorema non fa alcun riferimento alla presenza di norme su X; e quindi in particolare nessuna considerazione topologica può usarsi nella dimostrazione.

L'idea della dimostrazione del Teorema 119 è la seguente: si prova che se $Y \neq X$ allora esiste un'estensione lineare propria L_1 di L_0 e che L_1, definito sul sottospazio Y_1 propriamente contenente Y_0, soddisfa ancora alla disuguaglianza

$$L_1 x \leq p(x) \qquad \forall x \in Y_1 \,.$$

Per indicare che vale questa disuguaglianza, diremo che L_1 è *dominata* da p.

Questa è la parte tecnica della dimostrazione, che vedremo in seguito. Accettando ciò, si indichi con $\mathcal{F}$ la famiglia di tutte le estensioni di L_0 dominate da p; ossia, $L \in \mathcal{F}$ se L estende L_0 ed $Lx \leq p(x)$ per ogni x nel dominio di L.

In $\mathcal{F}$ si introduce una relazioni d'ordine parziale ponendo

$$L_1 \leq L_2 \qquad \text{se } L_2 \text{ estende } L_1.$$

Si vede immediatamente che ogni catena $\mathcal{C}$ in $\mathcal{F}$ ammette come maggiorante quell'elemento $\tilde{L} \in \mathcal{F}$ il cui dominio è l'unione dei domini di tutti gli elementi della catena $\mathcal{C}$ e così definito: se $x \in \mathrm{dom}\tilde{L}$ allora esiste $L' \in \mathcal{C}$ (non unico) il cui dominio contiene x. Si definisce

$$\tilde{L}x = L'x \,.$$

Si è detto che L' non è unico; ma l'essere $\mathcal{C}$ una catena prova che la trasformazione $\tilde{L}$ appena definita è univoca e lineare.

E' ovvio che $\tilde{L}$ è dominato da p.

Discende dal Lemma di Zorn che $\mathcal{F}$ ha un elemento massimale L: un'estensione L di L_0, dominata da p, non ulteriormente estendible, e quindi con dominio X.

Proviamo ora la parte tecnica della dimostrazione, consistente nella costruzione dell'estensione L_1. Prima però notiamo:

Osservazione 141 Sia X uno spazio di Banach di dimensione infinita e sia $\{e_n\}$ un sistema linearmente indipendente di elementi, tutti di norma 1. Definiamo

$$L_0 e_n = n$$

e quindi estendiamo L_0 a

$$\mathrm{span}\,\{e_n\} = \left\{ \sum_{\text{finita}} \alpha_i e_i \right\} .$$

L'argomento precedente basato sul *lemma di Zorn* può ripetersi anche senza far intervenire la funzione $p(x)$ e conduce a provare l'esistenza di un funzionale *L definito su X* e che estende L_0. Ciò mostra che esistono funzionali lineari *non continui, il cui dominio è tutto lo spazio X*. ∎

Veniamo ora a costruire una opportuna estensione L_1 di L_0, dominata da p.

Essendo $Y_0 \neq X$, esiste $x_0 \notin Y_0$. Scegliamo per Y_1 lo spazio lineare generato da Y_0 e da x_0,

$$Y_1 = \{y + tx_0 \mid y \in Y_0\,,\ t \in \mathbb{R}\} .$$

Se vogliamo che L_1 sia un'estensione lineare di L_0, dovrà essere

$$L_1(y + tx_0) = L_1 y + tL_1 x_0 = L_0 y + tL_1 x_0$$

e il problema si riduce a trovare un valore ξ per $L_1 x_0$, in modo tale che L_1 sia dominato da p. Si vuole cioè che valga

$$L_0 y + t\xi \le p(y + tx_0)\,, \qquad \forall y \in Y_0\,,\ t \in \mathbb{R} .$$

Distinguiamo i due casi $t > 0$ e $t < 0$. Se $t > 0$ allora si richiede

$$L_0 \frac{y}{t} + \xi \le p\left(\frac{y}{t} + x_0\right) \qquad \text{ossia} \qquad \xi \le p\left(\frac{y}{t} + x_0\right) - L_0 \frac{y}{t}\,; \tag{2.60}$$

Se $t < 0$ si nota che

$$p\,(y_0 + tx_0) = p\,(y_0 - |t|x_0) = |t|p\left(-\frac{y_0}{t} - x_0\right)$$

e quindi si richiede

si richiede

$$L_0 \frac{y}{t} + \xi \ge -\frac{1}{-t} p(y + tx_0) = -p(-\frac{y}{t} - x_0)\,,$$

ossia

$$\xi \ge -L_0 \frac{y_0}{t} - p(-\frac{y}{t} - x_0)\,. \tag{2.61}$$

Dunque è da provare l'esistenza di un numero ξ che verifica la (2.60) per $t > 0$ e la (2.61) per $t < 0$.

Notiamo che, essendo y arbitrario in Y_0, anche y/t, $t > 0$ e $-y/t$, $t < 0$ sono arbitrari elementi di Y_0. Dunque, le (2.60), (2.61) equivalgono a

$$\sup_{y\in Y_0}\{-L_0 y - p(-y - x_0)\} \le \xi \le \inf_{y\in Y_0}\{p(y + x_0) - L_0 y\}\,.$$

Dunque. il numero ξ esiste se si può provare che

$$-L_0\hat{y} - p(-\hat{y} - x_0) \le -L_0\tilde{y} + p(\tilde{y} + x_0) \qquad \forall \hat{y}\,,\ \tilde{y} \in Y_0\,.$$

Quest'eguaglianza equivale a

$$L_0(\tilde{y} - \hat{y}) \le p(\tilde{y} + x_0) + p(-\hat{y} - x_0)$$

e questa vale certamente perché, per ipotesi, L_0 è dominato da p. Dunque si ha:

$$L_0(\tilde{y} - \hat{y}) \le p(\tilde{y} - \hat{y}) = p(\tilde{y} + x_0 - x_0 - \hat{y}) \le p(\tilde{y} + x_0) + p(-x_0 - \hat{y})\,.$$

Ciò completa la dimostrazione del Teorema di Hahn-Banach nel caso in cui $\Phi = \mathbb{R}$. ∎

Dimostrazione del TEOREMA 120.

Il caso $\Phi = \mathbb{C}$ si fa discendere dal precedente, procedendo come segue.

Si nota che se X è uno spazio lineare complesso, allora esso è anche uno spazio lineare reale. Indichiamo allora con i simboli $X_{(R)}$ ed $Y_{0,(R)}$ gli spazi lineari *reali* i cui elementi sono quelli di X e di Y_0. Sia inoltre $L_{0,(R)}$ il funzionale lineare definito su $Y_{0,(R)}$ da

$$L_{0,(R)}x = \Re e\, L_0 x\,.$$

La (2.53) mostra che $L_{0,(R)}$ è dominato da p e quindi che esiste un'estensione $L_{(R)}$ di $L_{0,(R)}$ ad $X_{(R)}$, ancora dominata da p. Inoltre, essendo $L_R x \le p(x)$ per ogni $x \in X$, vale anche

$$L_{(R)}(-x) \le p(-x) = p(x)$$

ossia,

$$-p(-x) \le L_{(R)}(x) \le p(x)\,. \tag{2.62}$$

Definiamo ora l'operatore

$$Lx = L_{(R)}x - iL_{(R)}(ix)\,.$$

Si verifica immediatamente che quest'operatore è lineare su X (con $\Phi = \mathbb{C}$). Inoltre, se $x \in Y_0$, vale

$$Lx = \Re e\, L_0x - i\Re e\, [iL_0x] = \Re e\, L_0x - i\Re e\, \{i\Re e\, L_0x - \mathcal{I}m\, L_0x\} = L_0x\,.$$

Dunque, L estende L_0 e inoltre vale $|Lx| \le p(x)$ per ogni $x \in X$. Infatti se ciò non fosse potrebbe trovarsi x_0 tale che

$$|Lx_0| > p(x_0)\,.$$

Con $\theta = \arg Lx_0$ avremmo

$$L_Rx_0 = \Re e\, Lx_0 = |e^{i\theta}Lx_0| = |Lx_0| > p(x_0)\,.$$

Ciò contrasta con (2.62).

La dimostrazione è ora completa. ∎

Dimostrazione del TEOREMA 124.

Ricordiamo che si è supposto $\Phi = \mathbb{R}$.

Sia A convesso aperto e non vuoto. A meno di una traslazione, si può supporre $0 \in A$. Definiamo il seguente funzionale p_A su X:

$$p_A(x) = \inf\{t > 0 \mid x \in tA\}$$

dove con tA si intende l'insieme $\{ty \mid y \in A\}$.

E' chiaro che il funzionale p_A è non negativo, positivamente omogeneo. Inoltre, esso è subadditivo. Infatti, sia $\epsilon > 0$ e siano $t_1 = p_A(x) + \epsilon$, $t_2 = p_A(y) + \epsilon$. Vale quindi

$$(x + y) \in [p_A(x) + \epsilon]A + [p_A(y) + \epsilon]A = [p_A(x) + p_A(y) + 2\epsilon]A\,.$$

Dunque, per ogni $\epsilon > 0$ vale $p_A(x+y) \le p_A(x) + p_A(y) + 2\epsilon$. La subadditività segue dall'arbitrarietà di ϵ.

Consideriamo ora un punto $x_0 \notin A$ ed il s.spazio

$$X_0 = \{sx_0 \mid s \in \mathbb{R}\}\,.$$

La convessità di A implica la convessità di $A \cap X_0$ e quindi l'insieme $\{s \mid sx_0 \in A\}$ è un segmento a cui non appartiene 1. Si definisca il funzionale lineare L_0 su X_0,

$$L_0(sx) = s\,.$$

Questo funzionale verifica

$$L_0(x_0) = 1\,, \qquad L_0(sx_0) \le p_A(sx_0)\,.$$

La prima proprietà è immediata e la seconda è ovvia se $s \le 0$. Che essa valga anche se $s > 0$ si vede notando che $p_A(x_0) \ge 1$ e quindi

$$L_0(sx_0) = s \le sp_A(x_0) = p_A(sx_0)\,.$$

Notato ciò, il *Teorema di Hahn–Banach,* mostra che il funzionale L_0 si estende ad un funzionale L lineare e continuo su X, che ancora verifica $\langle\!\langle L, x\rangle\!\rangle \le p_A(x)$ e quindi

$$\langle\!\langle L, x\rangle\!\rangle \le p_A(x) \le 1 \quad \forall x \in A\,, \qquad \langle\!\langle L, x_0\rangle\!\rangle = 1\,. \ \blacksquare$$

Il funzionale p_A si chiama il FINZIONALE DI MINKOWSKI del convesso A.

Dimostrazione del TEOREMA 125.

Il Teorema 125 è immediata conseguenza del Teorema 124. Studiamo prima il caso in cui A è aperto. Fissiamo $a_0 \in A$ e $b_0 \in B$ e sia $x_0 = a_0 - b_0$. Sia inoltre

$$C = A - B = \{a - b \mid a \in A\,,\ b \in B\} = \bigcup_{b\in B}(A - b)\,.$$

E' facile vedere che C è aperto (anche se B non lo è); e che $x_0 \in C$. Dunque, $0 \in D = C - x_0$ mentre $x_0 \notin D$. Infatti, se fosse $x_0 \in D$, esisterebbero $\tilde a \in A$, $\tilde b \in B$ per cui

$$x_0 = \tilde a - \tilde b - x_0 \quad \text{ossia} \quad \tilde a = \tilde b$$

mentre per ipotesi A e B sono disgiunti.

Dunque, per il Teorema 124, x_0 si separa da D con un funzionale x^*:

$$\langle\!\langle x^*, d\rangle\!\rangle \le \langle\!\langle x^*, x_0\rangle\!\rangle \qquad \forall d \in D\,.$$

Dunque,

$$\langle\!\langle x^*, (a - b + x_0)\rangle\!\rangle < \langle\!\langle x^*, x_0\rangle\!\rangle \qquad \forall a \in A\,, \quad b \in B\,;$$

Ossia,

$$\langle\!\langle x^*, a\rangle\!\rangle < \langle\!\langle x^*, b\rangle\!\rangle\,.$$

Ciò prova la prima parte del teorema.

La dimostrazione della seconda parte si riconduce a quella della prima, grazie al lemma seguente:

Lemma 142 *Siano K e C due s.insiemi di X, K compatto non vuoto e C chiuso. Se $K \cap C = \emptyset$, allora esiste una sfera B di X tale che*

$$(K + B) \cap C = \emptyset\,.$$

<u>Dim.</u> Per ogni $k \in K$ esiste una opportuna sfera $B_{\epsilon(k)}$ tale che $k + 2B_{\epsilon(k)} \cap (C + B_{\epsilon(k)}) = \emptyset$. L'unione degli aperti $k_i + 2B_{\epsilon(k)}$ copre K e quindi esistono k_i, in numero finito, diciamo n, tali che

$$K \subseteq \bigcup_{i=1}^{n} \left(k_i + 2B_{\epsilon(k_i)}\right) .$$

Sia B la sfera di minimo raggio tra le sfere $B_{\epsilon(k_i)}$. Allora, $\bigcup \left(k_i + 2B_{\epsilon(k_i)}\right)$ non interseca $C + B$ e quindi vale anche:

$$K + B \subseteq \bigcup_{i=1}^{+\infty} \left(k_i + B_{\epsilon(k_i)} + B\right) \subseteq \bigcup_{i=1}^{n} \left(k_i + B_{\epsilon(k_i)} + B_{\epsilon(k_i)}\right)$$

e quest'insieme non interseca C. ∎

E' ora facile provare la seconda parte del Teorema 120. Sia B_ϵ una sfera tale che $A_1 = A + B_\epsilon$ non intersechi B. Gli insiemi A_1 e B verificano le ipotesi della prima parte del teorema e quindi esiste x^* tale che

$$\sup\{\langle\langle x^*, a\rangle\rangle \mid a \in A_1\} \le \inf\{\langle\langle x^*, b\rangle\rangle \mid b \in B\} .$$

Ciò prova che x^* separa A e B. Proviamo ora che la separazione è stretta. Notiamo che

$$\langle\langle x^*, A_1\rangle\rangle = \{\langle\langle x^*, x\rangle\rangle \mid x \in A_1\}$$

è un intervallo aperto di $\mathbb{R}$. Infatti, sia $\alpha = \langle\langle x^*, a_0\rangle\rangle \in \langle\langle x^*, A_1\rangle\rangle$ e $B_\sigma = \{x \mid ||x|| < \sigma\}$ una sfera tale che $a_0 + B_\sigma \in A_1$. Allora, $\langle\langle x^*, (a_0 + B_\sigma)\rangle\rangle$ è un intervallo non ridotto ad un sol punto perché il funzionale x^* non è zero; e dunque α è interno ad $\langle\langle x^*, A_1\rangle\rangle$. D'altra parte, $\langle\langle x^*, A\rangle\rangle$ è compatto e contenuto nell'aperto $\langle\langle x^*, A_1\rangle\rangle$ e quindi ha distanza positiva dall'insieme $\langle\langle x^*, B\rangle\rangle$, che non interseca $\langle\langle x^*, A_1\rangle\rangle$. Ciò prova che la separazione è stretta. ∎

Osservazione 143 Abbiamo enunciato i teoremi di separazione assumendo $\Phi = \mathbb{R}$ per semplicità. Tali teoremi possono estendersi al caso $\Phi = \mathbb{C}$. In questo caso le diseguaglianze valgono tra le parti reali. Si hanno cioè condizioni del tipo

$$\sup\{\Re e\, La \mid a \in A\} \le \inf\{\Re e\, Lb \mid b \in B\} .$$ ∎

2.10 Convergenza debole e debole stella

Si è visto che in uno spazio di Banach di dimensione infinita i compatti sono "rari". D'altra parte, si sa dai corsi precedenti che la proprietà di compattezza è cruciale nello studio dei problemi di minimo. Infatti:

Teorema 144 (di WEIERSTRASS) *Una funzione f* continua *su un insieme* compatto K *ammette sia minimo che massimo.*

Accenniamo alla dimostrazione, che dovrebbe essere nota: per provare l'esistenza del minimo, si costruisce una SUCCESSIONE MINIMIZZANTE (k_n) in K, ossia una successione tale che

$$\lim f(k_n) = \inf\{f(k) \mid k \in K\}\,.$$

Si usa la compattezza di K per estrarre una sottosuccessione (k_{n_r}) convergente a $k_0 \in K$; e la continuità di f per concludere

$$f(k_0) = \lim f(k_{n_r}) = \inf\{f(k) \mid k \in K\}\,. \tag{2.63}$$

Ciò prova che k_0 è punto di minimo.

Osservazione 145 Alla medesima conclusione si giunge se la funzione f, invece di essere continua, ha soltanto epigrafo chiuso, ossia è SEMICONTINUA INFERIORMENTE. Infatti in tal caso da $k_{n_r} \to k_0$, $f(k_{n_r}) \to \inf\{f(k) \mid k \in K\}$. Se l'epigrafo è chiuso si ha

$$(k_0, \inf\{f(k) \mid k \in K\}) \in \mathrm{Epi} f\,;$$

ossia, invece della (2.63), vale

$$f(k_0) \leq \lim f(k_{n_r}) = \inf\{f(k) \mid k \in K\}\,.$$

Ovviamente la disuguaglianza stretta non può valere, e quindi, nella sola ipotesi di semicontinuità inferiore, segue la (2.63); ossia segue che k_0 è punto di minimo. ∎

E' facile vedere che:

- se un insieme è compatto in una topologia, tale rimane anche in topologie *meno fini;*
- se una funzione è continua oppure semicontinua inferiormente rispetto ad una topologia, tale rimane anche in topologie *più fini.*

Dunque in uno spazio di Banach conviene introdurre topologie *meno fini* di quella della norma, in modo da avere più insiemi compatti; ma sufficientemente fini in modo tale che almeno opportune classi di funzioni continue rimangano, se non continue, almeno semicontinue inferiormente.

Noi ci limiteremo ad introdurre *concetti di convergenza* per topologie meno fini di quella della norma. Non descriveremo invece la topologia.

Definizione 146 Sia (x_n) una successione in uno spazio di Banach X. Diciamo che (x_n) CONVERGE DEBOLMENTE ad x_0 se

$$\lim \langle\langle x^*, x_n \rangle\rangle = \langle\langle x^*, x_0 \rangle\rangle \qquad \forall x^* \in X^* . \quad \blacksquare$$

Per indicare la convergenza debole si usa uno dei due simboli

$$\text{w}-\lim x_n = x_0 \qquad \text{oppure} \qquad x_n \rightharpoonup x_0 .$$

La definizione stessa di convergenza debole mostra che:

Teorema 147 *Ogni $x^* \in X^*$ è continuo rispetto alla convergenza debole di successioni di X.*

Supponiamo ora di sapere che stiamo lavorando nello spazio di Banach X^*, duale dello spazio di Banach X. In X^* si può definire la convergenza debole, come sopra, facendo intervenire il suo duale; ma si può anche definire la CONVERGENZA DEBOLE STELLA, come segue:

Definizione 148 Sia (x_n^*) una successione in X^*. Diciamo che (x_n^*) CONVERGE IN SENSO DEBOLE STELLA ad x_0^* se

$$\lim \langle\langle x_n^*, x \rangle\rangle = \langle\langle x_0^*, x \rangle\rangle \qquad \forall x \in X . \quad \blacksquare$$

Per indicare la convergenza debole stella si usa uno dei due simboli

$$\text{w}^*-\lim x_n^* = x_0^* \qquad \text{oppure} \qquad x_n^* \rightharpoonup x_0^* .$$

(la notazione con la mezza freccia è la stessa per la convergenza debole e la convergenza debole stella. Il contesto evita ambiguità).

E' facile vedere che *la convergenza in norma implica la convergenza debole, rispettivamente debole stella.* Il viceversa, in dimensione infinita, *non vale.*

Esempio 149 Sia $X = L^2(0, 2\pi)$. Vedremo che $(L^2(0, 2\pi))^*$ è isometricamente isomorfo ad $L^2(0, 2\pi)$ stesso. In particolare ogni $g \in L^2(0, 2\pi)$ si interpreta come elemento di $(L^2(0, 2\pi))^*$, quell'elemento definito da

$$f \longrightarrow \int_0^{2\pi} \bar{g}(x) f(x) \, \mathrm{d}x .$$

Si sa, dalla teoria della serie di Fourier, che ogni $f \in L^2(0, 2\pi)$ si rappresenta

$$f(x) = \sum_{k=-\infty}^{k=+\infty} f_k e^{ikx} , \qquad f_k = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) e^{-ikx} \, \mathrm{d}x = \langle\langle e_k, f \rangle\rangle$$

con $e_k = e^{ikx}$. Si sa, dalla teoria della serie di Fourier, che

$$\lim f_k = 0\,.$$

Questo vuol dire che,

$$\mathrm{w}^*\!-\lim e_k = 0\,.$$

Ma,

$$||e_k||_2 = \sqrt{2\pi}$$

e quindi la successione (e_k) *non* tende a zero in norma. ∎

Vale però:

Teorema 150 *Il teorema di unicità del limite vale sia per la convergenza debole che per la convergenza debole stella.*

Dim. Se

$$\mathrm{w}\!-\lim x_n = x_0 \qquad \text{ed anche} \qquad \mathrm{w}\!-\lim x_n = y_0$$

allora si ha

$$\langle\!\langle x^*, (x_0 - y_0)\rangle\!\rangle = \lim\{\langle\!\langle x^*, x_n\rangle\!\rangle - \langle\!\langle x^*, x_n\rangle\!\rangle\} = 0$$

per ogni $x^* \in X^*$. Dunque x_0 ed y_0 sono indistinguibili da elementi di X^*. *E' conseguenza del Teorema di Hahn-Banach che* $x_0 = y_0$.

L'asserto relativo alla convergenza debole stella è invece elementare: sia

$$\mathrm{w}^*\!-\lim x_n^* = x_0^* \qquad \text{ed anche} \qquad \mathrm{w}^*\!-\lim x_n^* = y_0^*\,.$$

Si ha

$$\langle\!\langle [x_0^* - y_0^*], x\rangle\!\rangle = \lim\{\langle\!\langle x_n^*, x\rangle\!\rangle - \langle\!\langle x_n^*, x\rangle\!\rangle\} = 0$$

per ogni $x \in X$. E quindi $x_0^* - y_0^*$ è il funzionale nullo, ossia $x_0^* = y_0^*$. ∎

Esistono alcune relazioni importanti tra la convergenza debole, oppure debole stella, e le proprietà che valgono in norma. Tra queste:

Teorema 151 *Ogni successione convergente in senso debole stella è limitata nella norma di* X^*.

Dim. Sia $x \in X$. Se $\mathrm{w}^*\!-\lim x_n^* = x^*$ allora

$$\lim\langle\!\langle x_n^*, x\rangle\!\rangle = \langle\!\langle x^*, x\rangle\!\rangle\,.$$

Dunque, per ogni $x \in X$ esiste un numero M_x, dipendente da x, tale che

$$|\langle\langle x_n^*, x\rangle\rangle| < M_x\,.$$

Il teorema di Banach-Steinhaus mostra che l'insieme dei funzionali x_n^* è limitato nella norma di X^*. ∎

Un asserto analogo vale anche per la convergenza debole, ed anzi il risultato relativo alla convergenza debole è un corollario del precedente. Per provarlo, abbiamo bisogno di una premessa. Ogni $x \in X$ si può vedere come funzionale lineare continuo su X^*. Basta associare ad esso il funzionale

$$x^* \to \langle\langle x^*, x\rangle\rangle\,. \tag{2.64}$$

Dunque, ogni elemento $x \in X$ si può porre in corrispondenza con un elemento $jx \in (X^*)^*$, duale di X^*.

Lo spazio $(X^*)^*$ si chiama il BIDUALE di X e si indica col simbolo X^{**}.

Vale:

Teorema 152 *La trasformazione j da X in X^{**} definita in (2.64) è isometrica.*

Dim. Dobbiamo provare che

$$||x||_X = ||jx||_{X^{**}}\,,$$

ossia:

$$||x||_X = \sup_{||x^*||_{X^*}} |\langle\langle x^*, x\rangle\rangle|\,.$$

Questo è l'asserto del Corollario (122). ∎

Osservazione 153 In particolare segue che j è una trasformazione *iniettiva.* Vedremo che, in generale, *essa non è suriettiva.* ∎

Possiamo ora provare:

Corollario 154 *Ogni successione debolmente convergente in X è limitata in norma.*

Dim. Sia (x_n) una successione in X, debolmente convergente ad x_0. Ciò vuol dire che

$$\lim\langle\langle x^*, x_n\rangle\rangle = \langle\langle x^*, x_0\rangle\rangle \qquad \forall x^* \in X^*\,,$$

ossia

$$\lim\langle\langle jx_n, x^*\rangle\rangle = \langle\langle jx_0, x^*\rangle\rangle \qquad \forall x^* \in X^* .$$

Dunque, la successione (jx_n) converge in X^{**} a jx_0, nel senso debole stella ed è quindi limitata in X^{**}. Il Teorema 152 implica che (x_n) è limitata in X. ∎

Osservando con attenzione la dimostrazione del Teorema 151 si vede che:

Teorema 155 *vale:*

- *se* $\mathrm{w}-\lim x_n = x_0$ *allora* $\liminf ||x_n||_X \geq ||x_0||_X$*;*
- *se* $\mathrm{w}^* - \lim x_n^* = x_0^*$ *allora* $\liminf ||x_n^*||_{X^*} \geq ||x_0^*||_{X^*}$*.*

Dim. Proviamo l'asserto relativo alla convergenza debole stella.

Per ogni x, con $||x||_X \leq 1$, vale

$$\langle\langle x_n^*, x\rangle\rangle \leq ||x_n^*||_{X^*}$$

e dunque

$$\langle\langle x_0^*, x\rangle\rangle = \lim\langle\langle x_n^*, x\rangle\rangle \leq \liminf ||x_n^*||_{X^*} .$$

Da qui:

$$||x_0^*||_{X^*} = \sup_{||x||_X=1} \langle\langle x_0^*, x\rangle\rangle = \sup_{||x||_X=1} \lim\langle\langle x_n^*, x\rangle\rangle \leq \liminf ||x_n^*||_{X^*} .$$

L'asserto relativo alla convergenza debole si deduce dall'asserto relativo alla convergenza debole stella mediante il Teorema 152:

$$||x_0||_X = \sup_{||x^*||_{X^*}=1} \langle\langle x^*, x_0\rangle\rangle = \sup_{||x^*||_{X^*}=1} \lim\langle\langle x^*, x_n\rangle\rangle = \sup_{||x^*||_{X^*}=1} \lim\langle\langle jx_n, x^*\rangle\rangle \leq$$
$$\sup_{||x^*||_{X^*}=1} \{\liminf ||jx_n||_{X^{**}} ||x^*||_{X^*}\} = \liminf ||jx_n||_{X^{**}} = \liminf ||x_n||_X . \ \blacksquare$$

Il Teorema 155 può riformularsi dicendo che *la norma di X è debolmente semicontinua inferiormente e la norma di X^* è debolmente stella semicontinua inferiormente.*

Grazie al *Teorema di Weierstrass,* segue che la norma ammette minimo sugli insiemi che sono compatti rispetto alla convergenza debole oppure debole stella.

Proviamo ora:

Teorema 156 *Sia A un insieme sequenzialmente debolmente chiuso in X. Allora, A è anche chiuso in norma. Se A è sequenzialmente debolmente stella chiuso in X^*, esso è anche chiuso nella norma di X^*.*

Dim. Sia A chiuso rispetto alla convergenza debole. Sia (x_n) una successione in A, che converge *in norma* ad un x_0.

La convergenza in norma implica la convergenza debole, e quindi (x_n) *converge debolmente* ad x_0. Dato che A è debolmente chiuso, si ha $x_0 \in A$; e quindi A è chiuso in norma. L'assero relativo alla convergenza debole stella si prova in modo analogo. ∎

L'esempio 149 mostra che la superficie di una sfera di $L^2(0, 2\pi)$, pur essendo chiusa in norma, non è debolmente chiusa. Combinando questo col teorema precedente si vede che che la topologia della convergenza debole è effettivamente *meno fine* di quella della norma.

Vale:

Teorema 157 *Sia A un sottoinsieme di uno spazio di Banach X. Sia A* convesso *e* chiuso *rispetto alla norma. L'insieme A è anche* chiuso *rispetto alla convergenza debole.*

Dim. Sia (x_n) una successione a valori in A, debolmente convergente ad x_0. Dobbiamo provare che x_0 appartiene ad A. Ciò discende dai teoremi di separazione: se $x_0 \notin A$, allora esiste $x^* \in X^*$ che separa x_0 da A. Inoltre, la separazione è stretta perché l'insieme A è chiuso in norma e l'insieme costituito dal solo punto x_0 è compatto. Dunque, esiste $\epsilon > 0$ tale che

$$\langle\langle x^*, x_0 \rangle\rangle + \epsilon < \inf \{ \langle\langle x^*, x \rangle\rangle \mid x \in A \} \leq \inf \{ \langle\langle x^*, x_n \rangle\rangle \} .$$

Ciò contrasta con l'ipotesi che (x_n) converge debolmente ad x_0. ∎

E' bene notare esplicitamente che l'asserto analogo per la topologia debole stella non vale. Però nella topologia debole stella successioni compatte si identificano facilmente:

Teorema 158 (di Alaoglu) *Ogni successione limitata in X^* ammette s.successioni convergenti in senso debole stella.*

La dimostrazione è posposta.

L'asserto analogo *non vale* in X e ciò suggerisce di dare un nome particolare agli spazi X tali che $jX = X^{**}$.

Definizione 159 Se uno spazio X è isometrico al suo biduale, lo spazio X si dice RIFLESSIVO.

In uno spazio riflessivo, la convergenza debole equivale alla debole stella e quindi:

Teorema 160 *Ogni successione limitata in uno spazio riflessivo ammette s.successioni debolmente convergenti.*

Il Teorema 157 ha un corollario che bene illustra la geometria della convergenza debole:

Corollario 161 (di MAZUR**)** *Sia (x_n) una successione debolmente convergente ad x_0. Esiste una successione (s_n) in X, tale che:*

- *$s_n = \sum_{k=1}^{n} \lambda_k x_k$ per opportuni numeri $\lambda_k \in [0,1]$ tali che $\sum_{k=1}^{n} \lambda_k = 1$;*
- *la successione (s_n) converge ad x_0 in norma.*

Ossia, x_0 è limite in norma di una successione ottenuta da combinazioni convesse degli x_n.

Dim. Sia $A = \overline{\mathrm{co}}\{x_n\}$, il più piccolo convesso chiuso contenente gli x_n. Per il Teorema 157, x_0 appartiene ad A. Da qui segue l'asserto. ∎

Applichiamo ora il Teorema 157 allo spazio $X \times \mathbb{R}$. Sia A l'epigrafo di una funzione convessa e continua su X. L'insieme A è convesso e chiuso in norma e quindi anche chiuso rispetto alla convergenza debole. Dunque vale:

Teorema 162 *Una funzione convessa e continua in norma è anche semicontinua inferiormente rispetto alla convergenza debole.*

Questo teorema ed il Teorema di Alaoglu mostrano che il Teorema di Weierstrass può applicarsi alla minimizzazione di funzionali convessi su insiemi limitati e chiusi su spazi riflessivi; e ciò è importantissimo nella teoria dell'ottimizzazione.

Per completare queste considerazioni, mostriamo che, in contrasto con l'asserto del Teorema di Alaoglu, in un generico spazio di Banach esistono successioni limitate che non ammettono s.successioni convergenti nella topologia debole; e quindi che *esistono spazi di Banach che non sono riflessivi.* Infatti:

Teorema 163 *Esistono successioni limitate in $L^1(0,1)$, prive di s.successioni debolmente convergenti; e quindi $L^1(0,1)$ non è riflessivo.*

Dim. Si è visto al Teorema 147 che ogni elemento $x^* \in X^*$ è continuo rispetto alla convergenza debole in X. Se ogni successione limitata avesse s.successioni debolmente convergenti, allora, per il Teorema di Bolzano-Weierstrass, ogni elemento di X^* raggiungerebbe massimo su

$$\left\{ x \mid \int_0^1 |x(s)| \, \mathrm{d}s \le 1 \right\} .$$

L'esempio riportato al Teorema 80 mostra che ciò non accade. ∎

Osservazione 164 Ciò mostra una profonda differenza tra lo spazio $L^1(a,b)$ e gli spazi $L^p(a,b)$ con $1 < p < +\infty$. Vedremo infatti che questi ultimi spazi sono riflessivi e quindi tutte le successioni limitate in $L^p(0,1)$, con $1 < p < +\infty$ ammettono s.successioni debolmente convergenti. ∎

Nel Capitolo **??** introdurremo una particolare classe di spazi di Banach, che si chiamano spazi di Hilbert. Proveremo, al Par. 3.7:

Teorema 165 *Ogni spazio di Hilbert è riflessivo.*

2.10.0.1 Continuità e continuità debole di operatori lineari

Fa parte della definizione di convergenza debole che ogni elemento di X^*, ossia ogni funzionale lineare su X che è continuo *in norma,* è anche continuo rispetto alla convergenza debole. E' importante sapere che gli elementi di X^* sono *tutti* i funzionali lineari su X, continui rispetto alla convergenza debole:

Teorema 166 *Sia ϕ un funzionale* lineare *su X. Esso è continuo rispetto alla convergenza debole se e solo se è continuo in norma.*

Dim. Se ϕ è continuo rispetto alla convergenza debole lo è anche in norma perché ogni successione convergente in norma converge anche debolmente, al medesimo limite. Il viceversa è provato al Teorema 147. ∎

Un asserto analogo vale in realtà per generici operatori lineari, ma la dimostrazione è più profonda:

Teorema 167 *Siano X ed Y due spazi di Banach e sia A un operatore lineare da X in Y, con dominio uguale ad X e debolmente continuo, ossia tale che*

$$\text{se} \qquad x_n \rightharpoonup x_0 \quad \text{in } X \qquad \text{allora} \qquad Ax_n \rightharpoonup Ax_0 \quad \text{in } Y.$$

In tal caso, l'operatore A è continuo in norma.

Dim. Proviamo che nelle ipotesi del teorema, l'operatore lineare A, definito su X, è chiuso.

Sia $x_n \to x_0$ tale che $Ax_n \to y_0$ (rispettivamente nelle *norme* di X e di Y). Allora, $x_n \rightharpoonup x_0$ e quindi, per le ipotesi, $Ax_n \rightharpoonup Ax_0$; D'altra parte, $Ax_n \to y$ implica $Ax_n \rightharpoonup y$. Per l'unicità del limite debole, $y = Ax_0$ ossia la successione di punti $((x_n, Ax_n))$ se converge, converge ad un punto del grafico, che quindi è chiuso.

Ciò prova che l'operatore lineare A è chiuso e quindi, avendo dominio uguale ad X, continuo. ∎

Si osservi che la dimostrazione precedente fa uso sia del teorema di Hahn-Banach che del Teorema di Baire.

2.10.1 Dimostrazioni posposte

Dimostrazione del TEOREMA 158.

Il teorema vale in qualsiasi spazio duale. Però la dimostrazione che presentiamo usa un'ipotesi ulteriore, non richiesta dal teorema: presentiamo la dimostrazione nel caso in cui lo spazio X è separabile. Dunque supponiamo l'esistenza di una successione (x_n) la cui immagine è densa in X. Se vale quest'ipotesi, la dimostrazione si ottiene usando il METODO DIAGONALE DI CANTOR.

Sia (y_n^*) una successione limitata in X^*, $||y_n^*|| < \alpha$. Vogliamo estrarne una s.successione convergente.

Consideriamo la successione numerica $(\langle\langle y_n^*, x_1\rangle\rangle)$. Questa è *limitata* e quindi ammette una s.successione convergente. Indichiamola $(\langle\langle y_{n,1}^*, x_1\rangle\rangle)$. Consideriamo quindi $(\langle\langle y_{n,1}^*, x_2\rangle\rangle)$. Anche questa successione è limitata e quindi se ne estrae una s.successione convergente, che indichiamo $(\langle\langle y_{n,2}^*, x_2\rangle\rangle)$. Procedendo in questo modo si costruisce la tavola seguente:

$\langle\langle y_{1,1}^*, x_1\rangle\rangle$	$\langle\langle y_{2,1}^*, x_1\rangle\rangle$	$\langle\langle y_{3,1}^*, x_1\rangle\rangle$	$\langle\langle y_{4,1}^*, x_1\rangle\rangle$	$\dots$
$\langle\langle y_{1,2}^*, x_2\rangle\rangle$	$\langle\langle y_{2,2}^*, x_2\rangle\rangle$	$\langle\langle y_{3,2}^*, x_2\rangle\rangle$	$\langle\langle y_{4,2}^*, x_2\rangle\rangle$	$\dots$
$\vdots$				
$\langle\langle y_{1,k}^*, x_k\rangle\rangle$	$\langle\langle y_{2,k}^*, x_k\rangle\rangle$	$\langle\langle y_{3,k}^*, x_k\rangle\rangle$	$\langle\langle y_{4,k}^*, x_k\rangle\rangle$	$\dots$
$\vdots$				

la successione $(y_{n,k}^*)$ che figura su ciascuna riga è s.successione di tutte le successioni $y_{n,r}^*$, con $r < k$; e quindi $(\langle\langle y_{n,k}^*, x_r\rangle\rangle)$ converge, per ogni $r \leq k$.

Dunque, la successione diagonale $(y_{n,n}^*)$ ha la proprietà che $(\langle\langle y_{n,n}^*, x_r\rangle\rangle)$ converge per ogni r.

Proviamo che in realtà la successione $(\langle\langle y_{n,n}^*, x_0\rangle\rangle)$ converge per ogni $x_0 \in X$. Basta provare che essa è fondamentale, dato che essa prende valori nel campo scalare.

Si fissi per questo $\epsilon > 0$ ed $x_0 \in X$, qualsiasi. Sia N tale che

$$||x_0 - x_N|| < \epsilon/(3\alpha) \qquad \text{(ricordiamo che } \|y_n^*\| < \alpha) \,.$$

Si noti che N dipende da x_0. Si stimi quindi

$$\begin{aligned} &|\langle\langle y_{m,m}^*, x_0\rangle\rangle - \langle\langle y_{n,n}^*, x_0\rangle\rangle| \\ &\leq |\langle\langle y_{m,m}^*, x_0 - x_N\rangle\rangle| + |\langle\langle y_{m,m}^* - y_{n,n}^*, x_N\rangle\rangle + |\langle\langle y_{n,n}^*, x_N - x_0\rangle\rangle| \,. \end{aligned}$$

Il primo e l'ultimo addendo sono minori di $\epsilon/3$, per la scelta fatta di x_N. L'addendo intermedio è minore di $\epsilon/3$ se m, n superano un numero $\tilde{N}$, che dipende da ϵ e da N ossia da ϵ e da x_0. Ciò prova la convergenza della successione di numeri $(\langle\langle y^*_{n,n}, x\rangle\rangle)$, per ogni $x \in X$.

Si costruisce ora un funzionale y^*_0 su X ponendo

$$\langle\langle y^*_0, x\rangle\rangle = \lim \langle\langle y^*_{n,n}, x\rangle\rangle \,.$$

E' immediato verificare che y^*_0 è lineare. Inoltre, y^*_0 è continuo perché la succesione (y^*_n) è limitata, $||y^*_n||_{X^*} < \alpha$ per ogni n così che per $||x||_X < 1$,

$$\left|\langle\langle y^*_0, x\rangle\rangle\right| = \lim \left|\langle\langle y^*_{n,n}, x\rangle\rangle\right| \leq \alpha \,.$$

Dunque y^*_0 è continuo. E quindi, $y^*_0 = \mathrm{w}^*\!-\lim y^*_{n,n}$. Ciò è quanto volevamo provare. ∎

Osservazione 168 Nel caso in cui il *duale* X^* di X sia separabile, può sembrare possibile applicare il ragionamento precedente ad una successione (x_n) di X, arrivando a provare un analogo del teorema di Alaoglu in X. Ci si convinca che *ciò è falso* esaminando bene la dimostrazione e notando che, tentando di ripetere la dimostrazione, *NON* si proverebbe la convergenza debole di (x_n) in X, ma la convergenza *debole stella* di (jx_n) in X^{**}. ∎

2.10.2 Relazione tra le convergenze debole e debole stella

Avendo a disposizione gli esempi precedenti, possiamo chiarire meglio le relazioni tra le convergenze debole e debole stella, quando queste si possano definire sul medesimo spazio, che in tal caso è lo spazio X^* duale di uno spazio di Banach X.

Queste due nozioni di convergenza non sono indipendenti. Infatti:

Teorema 169 *Sia x^*_n una successione in X^*. Se essa converge debolmente ad x^*_0, allora essa converge anche debole stella al medesimo x^*_0.*

Dim. Ciò discende dal fatto che si è già notato che X è isometrico ad un s.spazio di X^{**}. ∎

L'esempio seguente mostra che non vale l'implicazione inversa.

Esempio 170 Sia $X = c_0$, $X^* = l^1$ ed $X^{**} = l^\infty$.

In $X^* = l^1$ si consideri la successione $\{e^{(n)}\}$ definiti da (2.31).

Se $x \in c_0$, $x = (x_n)$ allora

$$\langle\langle x, e^{(n)} \rangle\rangle = x_n \to 0 \qquad \text{e quindi} \qquad \text{w}^*\!-\lim e^{(n)} = 0\,.$$

Invece, se $x \in l^\infty = X^{**}$ è la successione ogni cui elemento vale 1,

$\langle\langle x, e^{(n)} \rangle\rangle = 1$ e quindi la successione $(e^{(n)})$ *non* converge debolmente a 0. ∎

2.11 Inversi di un operatore

In dimensione finita, l'equazione

$$Ax = y\,, \qquad x\,,\ y \in \mathbb{C}^n$$

con x, y vettori ed A trasformazione lineare, è risolubile per ogni y se e solo se

$$\ker A = \{0\}$$

ed in tal caso esiste l'operatore inverso A^{-1} di A che è lineare e che verifica ambedue le condizioni

$$\begin{cases} AA^{-1} &= I \\ A^{-1}A &= I\,; \end{cases}$$

anzi, se un operatore indicato con A^{-1} soddisfa una delle due uguaglianze precedenti esso soddisfa anche la seconda ed è l'operatore inverso di A, si veda il paragrafo (1.1.2)[19]. La situazione è più complessa in dimensione infinita. Infatti:

Esempio 171 Sia $X = l^p$, per un qualsiasi $p \in [1, +\infty]$ e siano T ed S definiti da

$$\begin{aligned} S\left[\begin{array}{cccc} x_0 & x_1 & x_2 & \dots \end{array}\right] &= \left[\begin{array}{cccc} 0 & x_0 & x_1 & \dots \end{array}\right] \\ T\left[\begin{array}{cccc} x_0 & x_1 & x_2 & \dots \end{array}\right] &= \left[\begin{array}{cccc} x_1 & x_2 & x_3 & \dots \end{array}\right]. \end{aligned} \tag{2.65}$$

Questi operatori si chiamano le TRASLAZIONI rispettivamente destra e sinistra, su l^p.

Sia ora $X = Y = l^2$ e consideriamo le traslazioni S e T su l^2. Questi operatori permettono di mostrare:

- Un operatore può *avere nucleo nullo senza essere suriettivo.* L'operatore S ha questa proprietà.
- Un operatore *può avere nucleo non nullo ed essere suriettivo.* Un operatore con tale proprietà è l'operatore T.
- Se un operatore lineare A da X in sé ha nucleo nullo allora esiste B definito su $\operatorname{im} A$ e tale che $BAx = x$ per ogni $x \in \operatorname{dom} A$. *Può accadere che non valga $ABx = x$.* Un operatore A con tale proprietà è $A = S$ (si scelga per B la restrizione di T ad $\operatorname{im} S$).

[19] si ricordi che x ed y appartengono ambedue a $\mathbb{C}^n$. Se essi appartengono a spazi diversi allora l'ultima affermazione è falsa.

- Si verifica facilmente l'esistenza di un operatore lineare B che verifica $TBx = x$ per ogni $x \in l^2$. Si scelga infatti $B = S$. Non vale però $BTx = x$ per ogni $x \in l^2$. ∎

Queste considerazioni suggeriscono la seguente definizione:

Definizione 172 Sia K un operatore lineare, limitato o meno, da X in Y. Se un *operatore* A, con dominio $\operatorname{im} K$, verifica

$$AKx = x \qquad \forall x \in X\,,$$

allora l'operatore A si chiama INVERSO SINISTRO di K. Se un *operatore lineare* B verifica

$$KBy = y \qquad \forall y \in Y$$

allora l'operatore B si chiama INVERSO DESTRO di K.

Un operatore che è sia inverso destro che sinistro di K si chiama INVERSO di K e si indica col simbolo K^{-1}. ∎

Si noti che la definizione di linearità è stata esplicitamente richiesta nella definizione di inverso destro, ma non in quella di inverso sinistro. Ciò perché:

Teorema 173 *Un operatore lineare K da X in Y ammette inverso sinistro se e solo se* $\ker K = \{0\}$*. In tal caso l'inverso sinistro è unico, ed è lineare.*

Dim. Se esiste l'inverso sinistro A di K allora per ogni $x \in \operatorname{dom} K$ vale:

$$x = AKx$$

e quindi se $Kx = 0$ si ha anche $x = 0$. Dunque, l'esistenza dell'inverso sinistro implica $\ker K = \{0\}$, ossia che K è iniettivo.

Viceversa, sia $\ker K = \{0\}$. Allora K, essendo lineare, è (univoco e) iniettivo. Il suo inverso sinistro è l'operatore che a Kx associa x.

La definizione di inverso sinistro mostra che il suo grafico in $Y \times X$ è l'insieme $\{(Kx, x) \mid x \in X\}$. Questo è un s.spazio perché l'operatore K è lineare; e quindi anche l'inverso sinistro, avendo per grafico un s.spazio, è lineare, si veda il teorema 65. ∎

Per contrasto si noti che *l'inverso destro non è unico* e che possono anche esistere operatori B non lineari che verificano l'uguaglianza $KBy = y$ per ogni y:

Esempio 174 Sia $X = Y = l^2$. Sia T l'operatore introdotto nell'esempio 171. Per ogni numero naturale n e per ogni numero reale α, definiamo $B_{n,\alpha}$ ponendo

$$B_{n,\alpha}(y_1, y_2, y_3, \dots) = (\alpha y_1^n, y_1, y_2, y_3, \dots)\,.$$

L'operatore $B_{n,\alpha}$ è non lineare se $n > 1$ ed è lineare se $n = 1$. Per ogni scelta di n e di α si ha: $TB_{n,\alpha}y = y$ per ogni $y \in l^2$.

In particolare si vede la non unicità dell'inverso destro perfino con la condizione che esso debba essere lineare. ∎

Le considerazioni precedenti suggeriscono di privilegiare lo studio dell'inverso sinistro. Approfondendo tale studio, notiamo che non c'è relazione tra continuità di un operatore e continuità del suo inverso sinistro, come ora vediamo:

Esempio 175 Mostriamo l'esempio di un operatore continuo ed invertibile, con inverso sinistro non continuo.

Sia $X = Y = l^2$ e sia A l'operatore definito da

$$A(x_1, x_2, x_3, x_4, \dots) = (x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \frac{1}{4}x_4, \dots)\,.$$

E' facile vedere che l'operatore A *è continuo e iniettivo,* e che il suo inverso sinistro, definito su $\operatorname{im} A$, è l'operatore *non continuo* B:

$$B(y_1, y_2, y_3, \dots) = (y_1, 2y_2, 3y_3, \dots)\,.$$

Per avere un esempio di operatore illimitato il cui inverso è limitato, si scambino i ruoli degli operatori A e B appena introdotti. ∎

E' facile dare un test per la limitatezza dell'inverso sinistro: sia B inverso sinistro di A. Per definizione, l'operatore B è limitato se e solo se esiste $\rho > 0$ tale che

$$||By||_X \le \rho ||y||_Y \qquad \forall y \in \operatorname{dom} B\,.$$

Un elemento y è in $\operatorname{dom} B$ se e solo se esiste x per cui $y = Ax$ e quindi la disuguaglianza precedente equivale a

$$m||x||_X \le ||Ax||_Y \tag{2.66}$$

(con $m = 1/\rho > 0$). Questa condizione implica anche che $\ker A = \{0\}$. Dunque:

Teorema 176 *L'operatore A ammette inverso sinistro continuo se e solo se esiste $m > 0$* (disuguaglianza stretta!) *per cui vale (2.66).*

La condizione (2.66) non è di facile verifica ed in generale non è facile costruire l'espressione esplicita dell'inverso. Un caso semplice ed importante è il seguente:

Teorema 177 *Sia $A \in \mathcal{L}(X)$, con $||A|| < 1$* (disuguaglianza stretta!) *e si consideri l'operatore $I - A$. L'operatore $I - A$ è iniettivo e suriettivo, ossia invertibile, ed è*

$$(I - A)^{-1} = \sum_{k=0}^{+\infty} A^k \in \mathcal{L}(X) \,. \tag{2.67}$$

<u>Dim</u>. Sia $q < 1$ tale che $||A|| < q$, ossia tale che $||Ax|| \leq q||x||$ per ogni x. Vale:

$$||(I - A)x|| = ||x - Ax|| \geq \Big| ||x|| - ||Ax|| \Big| = ||x|| - ||Ax|| \geq (1 - q)||x|| \,.$$

Dunque, l'operatore $I - A$ ammette l'inverso sinistro continuo, per il Teorema 176. Per trovare un'espressione per l'inverso sinistro consideriamo la serie in (2.67) (suggerita dalla serie geometrica!) Mostriamo prima di tutto che essa converge in $\mathcal{L}(X)$. Per questo consideriamo la successione delle somme parziali

$$S_n = \sum_{k=1}^{n} A^k \,.$$

Si ha:

$$||S_n - S_{n+m}|| = \left|\left| \sum_{k=n+1}^{n+m} A^k \right|\right| \leq \sum_{k=n+1}^{n+m} ||A^k|| \leq \sum_{k=n+1}^{n+m} q^k \,.$$

Essendo $q \in [0, 1)$ si ha che la successione delle somme parziali è fondamentale e quindi convergente in $\mathcal{L}(X)$.

Notiamo che, in particolare,

$$||A^k|| \leq ||A||^k \leq q^k \qquad \text{così che} \quad \lim_k A^k = 0 \,.$$

Per definizione

$$\left(\sum_{k=1}^{+\infty} A^k \right) x = \left(\lim_k \sum_{k=1}^{n} A^k \right) x = \lim_k \left(\sum_{k=1}^{n} A^k x \right) = \sum_{k=1}^{+\infty} A^k x \,.$$

Dunque:

$$\left(\sum_{k=0}^{+\infty} A^k\right)(I-A)x = \sum_{k=0}^{+\infty} A^k x - \sum_{k=0}^{+\infty} A^{k+1}x$$
$$= \lim\left\{\sum_{k=0}^{n} A^k x - \sum_{k=0}^{n} A^{k+1}x\right\} = \lim\{x - A^{k+1}x\} = x$$

(perché $\|A^k\| \to 0$) e quindi la serie in (2.67) rappresenta l'inverso sinistro di $(I-A)$. Con calcoli analoghi si vede che è anche inverso destro, e quindi inverso. In particolare segue che $I-A$ è suriettivo. ∎

La serie (2.67) si chiama SERIE DI VON NEUMANN.

Sottolineiamo ora che, per definizione, un operatore ammette inverso quando ammette sia inverso destro che sinistro; in particolare quando è sia iniettivo che suriettivo. Conviene indebolire un po' questa definizione.

Definizione 178 Sia A lineare da X in Y *con dominio denso in* X *e con immagine densa in* Y. Sia A iniettivo. L'inverso sinistro di A, definito su $\operatorname{im} A$ ed a valori in $\operatorname{dom} A$, si chiama INVERSO di A.

Il simbolo A^{-1} si usa anche per indicare l'inverso di A, nel senso generalizzato che abbiamo ora definito. ∎

Notiamo infine:

Teorema 179 *Si ha:*

- *L'inverso A^{-1}, nella versione debole della definizione 178, se esiste è lineare.*
- *Siano A e B operatori lineari ambedue invertibili con*

$$\operatorname{im} B \subseteq \operatorname{dom} A\,.$$

Allora $(AB)^{-1}$ esiste e vale

$$(AB)^{-1} = B^{-1}A^{-1}\,.$$

Dim. La prima affermazione vale perché l'inverso, secondo la definizione 178, è l'inverso sinistro. La seconda affermazione è immediata notando che $\operatorname{dom} AB = \operatorname{dom} B$ e che

$$B^{-1}A^{-1}ABx = x \qquad \forall x \in \operatorname{dom} A\,. \ ∎$$

2.12 Lo spettro di un operatore

> Per leggere questo paragrafo è necessario conoscere l'enunciato del Teorema 121: *un elemento x di uno spazio di Banach X è non nullo se e solo se esiste $x^* \in X^*$ tale che $\langle\langle x^*, x\rangle\rangle \neq 0$.* Si esprime a parole questa proprietà dicendo che LO SPAZIO DUALE DISTINGUE i punti di X.

E' noto che molte informazioni su un operatore lineare che opera su uno spazio lineare di dimensione finita si ottengono studiandne gli autovalori, i quali hanno spesso interpretazioni fisiche importanti. Vogliamo ora estendere questo tipo di studio a generici spazi di Banach.

Si sa che, anche in dimensione finita, autovalori ed autovettori possono trovarsi solo se il campo scalare è quello dei numeri complessi. Per questo supporremo di lavorare con spazi lineari su $\mathbb{C}$.

In dimensione finita, un numero complesso z_0 si dice un autovalore di A se

$$(z_0 I - A)x = y$$

non è risolubile per ogni y; equivalentemente, se la soluzione x, quando esiste, *non* è unica. L'esempio seguente mostra che la non risolubilità per ogni y in dimensione infinita *non* equivale alla non unicità. Ciò si è già visto all'Esempio 171. E' un fatto così importante che conviene ripeterlo.

Esempio 180 Consideriamo ancora gli operatori di traslazione S e T (definiti in (2.65)) su l^2. E' chiaro che

$$\begin{aligned} \ker S = \{0\} &\qquad \text{e} \quad \operatorname{im} S \neq X\,, \\ \ker T \neq \{0\} &\qquad \text{e} \quad \operatorname{im} T = X\,. \ \blacksquare \end{aligned}$$

Ricordiamo inoltre che se l'equazione

$$(zI - A)x = y \tag{2.68}$$

è risolubile in modo unico allora $(zI - A)$ ammette inverso (ossia inverso sinistro, si veda la definizione 178), definito sulla sua immagine e questo è un operatore lineare. Ma, in generale, l'inverso non è continuo.

Queste considerazioni suggeriscono la definizione seguente, che si applica ad ogni operatore lineare A da X in X, *anche non continuo ma con dominio denso:*

Definizione 181 Sia A lineare da X in X, con dominio denso. Si chiama INSIEME RISOLVENTE di A l'insieme dei numeri complessi z per cui la (2.68) ammette un'*unica* soluzione x per *ogni* y in un s.insieme *denso* di X e inoltre l'inverso $(zI - A)^{-1}$ è *continuo.*

Se z appartiene all'insieme risolvente di A, l'operatore $(zI - A)^{-1}$ si chiama l'OPERATORE RISOLVENTE di A.

L'insieme risolvente si indica col simbolo $\rho(A)$. Il suo complementare si indica col simbolo $\sigma(A)$ e si chiama lo SPETTRO dell'operatore A. ∎

Da un punto di vista logico, $z \in \sigma(A)$ se si verifica uno dei casi seguenti, mutuamente incompatibili:

i) $\ker(zI - A) \neq \{0\}$;

ii) $\ker(zI - A) = \{0\}$ ma $\mathrm{im}\,(zI - A)$ *non denso* in X;

iii) $\ker(zI - A) = \{0\}$, $\mathrm{im}\,(zI - A)$ denso in X ma $(zI - A)^{-1}$ *non continuo.*

Definiamo quindi:

- SPETTRO DI PUNTI l'insieme dei numeri z per i quali si verifica il caso *i)*;
- SPETTRO RESIDUO l'insieme dei punti per i quali si verifica il caso *ii)*;
- SPETTRO CONTINUO l'insieme dei punti per i quali si verifica il caso *iii).*

Gli elementi dello spettro di punti si chiamano AUTOVALORI dell'operatore A. Se λ è un autovalore, l'insieme

$$\{x \mid Ax = \lambda x\}$$

è un sottospazio che si chiama L'AUTOSPAZIO dell'autovalore λ.

Abbiamo definito una partizione dello spettro di A in tre s.insiemi. Essi si indicano rispettivamente con i simboli

$$\sigma_p(A)\,, \qquad \sigma_r(A)\,, \qquad \sigma_c(A)\,.$$

In dimensione finita solo il caso *i)* può verificarsi. Mostriamo che, invece, in dimensione infinita anche gli altri casi possono verificarsi.

Esempio 182 Siano S e T gli operatori di traslazione definiti in (2.65). Si vede facilmente che $0 \in \sigma_r(S)$. Invece, $0 \in \sigma_p(T)$.

Mostriamo un operatore con spettro continuo non vuoto. Sia $X = l^2$ e sia A definito come segue:

$$A \left[\begin{array}{cccccc} x_0 & x_1 & x_2 & \dots & x_n & \dots \end{array} \right] = \left[\begin{array}{cccccc} x_0 & \frac{1}{2}x_1 & \frac{1}{3}x_2 & \dots & \frac{1}{n+1}x_n & \dots \end{array} \right] .$$

L'equazione $Ax = y$ è risolubile per ogni successione (y_n) definitivamente nulla, ossia per ogni y in un s.spazio denso di $X = l^2$, ed è

$$\left[\begin{array}{cccccc} x_0 & x_1 & x_2 & \dots & x_n & \dots \end{array} \right] = \left[\begin{array}{cccccc} y_0 & 2y_1 & 3y_2 & \dots & (n+1)y_n & \dots \end{array} \right] .$$

Dunque, l'inverso non è continuo. ∎

Grazie al teorema fondamentale dell'algebra, si sa che in dimensione finita lo spettro non è mai vuoto ed è un insieme finito. Mostriamo invece che esistono operatori lineari su spazi di Banach, con spettro vuoto ed operatori con risolvente vuoto.

Esempio 183 Sia $X = L^2(0,1)$ e siano A e B definiti come segue:

$$\operatorname{dom} A = \{x \in C(0,1) \mid x' \in L^2(0,1)\} , \qquad Ax = x' ,$$
$$\operatorname{dom} B = \{x \in C(0,1) \mid x' \in L^2(0,1) , \ x(0) = 0\} , \qquad Bx = x' .$$

Allora $\sigma(A) = \sigma_p(A) = \mathbb{C}$ perché per ogni z vale $(z-A)\chi_z = 0$, con $\chi_z(t) = e^{zt}$. Invece, $\sigma(B) = \emptyset$ perché

$$(zI - B)x = y \Longleftrightarrow \left\{ \begin{array}{l} x' = zx - y \\ x(0) = 0 \, . \end{array} \right.$$

Dunque x è dato da

$$x(t) = -\int_0^t e^{z(t-s)} y(s) \, \mathrm{d}s$$

così che l'operatore $(zI - B)^{-1}$ è continuo per ogni z e quindi $\sigma(B) = \emptyset$. ∎

Nell'esempio precedente intervengono, vedremo non per caso, operatori che non sono continui; ma anche lo spettro di operatori continui può avere una struttura piuttosto complessa:

Esempio 184 Sia $X = l^2$ e sia T l'operatore definito in (2.65). Risolvendo

$$(zI - T)x = 0$$

si trova come soluzione:

$$x = q \left[\begin{array}{ccccc} 1 & z & z^2 & z^3 & \dots \end{array} \right]$$

con $q \in \mathbb{C}$ qualsiasi. Questa successione appartiene ad l^2 per ogni z di modulo minore di 1. Dunque, $\sigma_p(T) \supseteq \{z \mid |z| < 1\}$. ∎

Nonostante questi esempi, spettro e risolvente non possono essere insiemi qualsiasi. Infatti:

Teorema 185 *Se A è continuo, $\sigma(A) \subseteq \{z \mid |z| \leq ||A||\}$.*

Dim. Supponiamo che A sia continuo e che sia $|z| > ||A||$. Vogliamo provare che in tal caso $z \in \rho(A)$. Scriviamo per questo

$$(zI - A) = z(I - K)\,, \qquad K = \frac{1}{z}A \quad \text{così che} \quad ||K|| < 1\,.$$

Dunque,

$$(zI - A)^{-1} = \frac{1}{z}\sum_{n=0}^{+\infty} K^n = \frac{1}{z}\sum_{n=0}^{+\infty} \frac{A^n}{z^n}\,. \tag{2.69}$$

Quest'operatore è continuo, dato che $||K|| = ||A/z|| < 1$; e quindi $z \in \rho(A)$. ∎

Lo spettro non può essere un insieme qualsiasi nemmeno se l'operatore A non è continuo. Infatti:

Teorema 186 *Il risolvente è sempre un insieme aperto e quindi lo spettro è chiuso.*

Dim. Sia A qualsiasi, anche non continuo. Proviamo che il suo risolvente è aperto. Se esso è vuoto niente va provato. Dunque supponiamo che esista un numero $z_0 \in \rho(A)$ e mostriamo che esso è *interno* al risolvente; ossia proviamo l'esistenza di $\epsilon > 0$ (che dipende sia da z_0 che da A) tale che se $|z| < \epsilon$ allora $z + z_0 \in \rho(A)$. Per questo scriviamo

$$(z + z_0)I - A = zI + (z_0 I - A) = (z_0 I - A)\left[I + z(z_0 I - A)^{-1}\right]\,. \tag{2.70}$$

Per il Teorema 185, l'operatore

$$\left[I + z(z_0 I - A)^{-1}\right]$$

è invertibile se

$$|z| < \epsilon = \frac{1}{||(z_0 I - A)^{-1}||}$$

e in tal caso $(z + z_0)I - A$ è invertibile con inverso limitato, essendo composizione di operatori invertibili ciascuno con inverso limitato. ∎

La (2.70) permette anche di trovare un'espressione per $[(z + z_0)I - A]^{-1}$:

$$\begin{aligned} [(z + z_0)I - A]^{-1} &= \left\{\sum_{k=0}^{+\infty} z^n \left[(z_0 I - A)^{-1}\right]^n\right\}(z_0 I - A)^{-1} \\ &= \sum_{k=0}^{+\infty} z^n \left[(z_0 I - A)^{-1}\right]^{n+1}\,. \end{aligned} \tag{2.71}$$

Dunque, fissato $z_0 \in \rho(A)$, la funzione $[(z+z_0)I - A]^{-1}$ si esprime come serie di potenze di z, a coefficienti operatori limitati.

Chiameremo FUNZIONI OLOMORFE a valori operatori quelle funzioni di $z \in \mathbb{C}$ che si esprimono localmente, in un opportuno intorno di ogni punto z_0 del loro dominio, mediante serie

$$\sum_{n=0}^{+\infty} K^n (z - z_0)^n$$

convergenti (nella norma di $\mathcal{L}(X)$). Dunque:

Corollario 187 *Se l'operatore A ha risolvente non vuoto, la funzione $z \to (zI - A)^{-1}$ è olomorfa su $\rho(A)$.*

Osservazione 188 Combinando il calcolo dell'esempio 184 con i teoremi 185 e 186, si vede che

$$\sigma(T) = \{z \mid |z| \le 1\} . \blacksquare$$

Torniamo ora a considerare un *operatore continuo* A. Si è detto che il suo risolvente non è vuoto, e anzi contiene l'esterno del disco $\{z \mid |z| \le ||A||\}$. Naturalmente, esso può anche estendersi all'interno di tale disco; ma non può riempirlo. Infatti:

Teorema 189 *Sia $A \in \mathcal{L}(X)$. Lo spettro di A non è vuoto.*

Dim. La dimostrazione procede per assurdo, e va confrontata con la dimostrazione del teorema fondamentale dell'algebra.

Dal Corollario 187 si sa che, se $\sigma(A) = \emptyset$, la funzione $(zI - A)^{-1}$ è olomorfa su $\mathbb{C}$. Dunque, per ogni $x \in X$, $y^* \in X^*$, la funzione

$$z \to f(z) = \langle\langle y^*, (zI - A)^{-1}x \rangle\rangle$$

è una funzione intera.

Dalla (2.71) si vede che per $|z| \to +\infty$, la funzione $f(z)$ tende a zero; e quindi $f(z)$ è una funzione intera e limitata, e quindi costante. Dato che un suo limite è nullo, essa deve essere identicamente zero.

Dunque, abbiamo provato che

$$\langle\langle y^*, (zI - A)^{-1}x \rangle\rangle \equiv 0 \qquad \forall x \in X\,, \quad \forall y^* \in X^* .$$

Dato che X^* distingue punti di X, deve essere

$$(zI - A)^{-1}x \equiv 0$$

per ogni $x \in X$. Ciò non può darsi perché $(zI - A)(zI - A)^{-1}x = x$ per ogni $x \in X$.

La contraddizione trovata prova il teorema. ∎

Nonostante che lo spettro di un operatore continuo non possa essere vuoto, può essere che esso sia un insieme molto più piccolo del disco di raggio $||A||$. Per esempio, in dimensione 2, la trasformazione lineare la cui matrice è

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

ha norma 1 ed il solo autovalore 0. E' un utile esercizio vedere che un caso analogo può darsi anche in dimensione infinita.

Esempio 190 Sia $X = L^2(0,1)$ e sia

$$(Ax)(t) = \int_0^t x(s)\,\mathrm{d}s\,.$$

Si sa che A è continuo e si vede immediatamente che $0 \in \sigma(A)$, dato che A^{-1} è l'operatore di derivazione, $A^{-1}y = y'$, che non è continuo.

Mostriamo che ogni altro numero z appartiene al risolvente. Per questo risolviamo

$$(zI - A)x = y \quad \text{ossia} \quad zx(t) - \int_0^t x(s)\,\mathrm{d}s = y(t)\,.$$

Dividendo per z si trova

$$x(t) - \frac{1}{z}y(t) = \frac{1}{z}\int_0^t x(s)\,\mathrm{d}s = \frac{1}{z}\int_0^t \left[x(s) - \frac{1}{z}y(s)\right]\,\mathrm{d}s + \int_0^t \frac{1}{z^2}y(s)\,\mathrm{d}s\,.$$

Quest'uguaglianza mostra che la funzione

$$\xi(t) = x(t) - \frac{1}{z}y(t)$$

è derivabile quasi ovunque, che $\xi(0) = 0$ e che

$$\xi'(t) = \frac{1}{z}\xi(t) - \frac{1}{z^2}y(t) \qquad \text{ossia} \qquad \xi(t) = -\frac{1}{z^2}\int_0^t e^{\frac{1}{z}(t-s)}y(s)\,\mathrm{d}s\,.$$

Da qui,

$$x(t) = \frac{1}{z}y(t) - \frac{1}{z^2}\int_0^t e^{\frac{1}{z}(t-s)}y(s)\,\mathrm{d}s\,.$$

La trasformazione da y ad x è, per ogni fissato $z \neq 0$, lineare e continua. Dunque, $\sigma(A) = \{0\}$. ∎

Questi esempi suggeriscono di chiamare RAGGIO SPETTRALE $r(A)$ il numero

$$r(A) = \max\{|z| \mid z \in \sigma(A)\}\,.$$

Il raggio spettrale di un operatore continuo si esprime in modo che richiama la formula per il raggio di convergenza di una serie di potenze:

Teorema 191 *Sia $A \in \mathcal{L}(X)$. Vale:*

$$r(A) = \lim \sqrt[n]{||A^n||}\,.$$

Dim. Si prova, esattamente come nel caso scalare, che una serie di potenze a valori operatori converge in un disco di centro z_0, che si chiama ancora DISCO DI CONVERGENZA,. Questo disco coincide col disco di convergenza della serie di potenze

$$\sum_{n=0}^{+\infty} ||A^n||(z - z_0)^n\,.$$

Applicando questo alla serie (2.69) si vede che

$$r(A) = \limsup \sqrt[n]{||A^n||}\,.$$

Si deve ora provare che in realtà esiste

$$\lim \sqrt[n]{||A^n||}\,.$$

Risulta più semplice provare l'esistenza del limite

$$\lim \frac{1}{n} \log ||A^n||\,.$$

Notiamo che

$$\log ||A^{n+m}|| \le \log ||A^n|| \cdot ||A^m|| \le \log ||A^n|| + \log ||A^m||\,.$$

Una successione di numeri (a_n), tutti positivi, che verificano

$$a_{n+m} \le a_n + a_m$$

si dice SUBADDITIVA. La dimostrazione del Teorema 191 si completa usando il lemma seguente:

Lemma 192 *Se la successione (a_n) è subadditiva allora esiste*

$$\lim \frac{1}{n} a_n\,.$$

Dim. Si fissa un numero naturale m e si studiano i quozienti a_n/n con $n > m$. Notiamo che si può scrivere

$$n = md + r\,, \qquad 0 \le r < m$$

e quindi, usando la subadditività,

$$a_n = a_{md+r} \le da_m + a_r\,.$$

Dividiamo per n e passiamo al limite per $n \to +\infty$.

Il numero a_r è funzione di n limitata al variare di n perché prende valori nell'insieme finito $a_1, \ldots, a_{m-1}$. Dunque $\lim a_r/n = 0$.

Ancora perché r prende un numero finito di valori,

$$\frac{d}{n} = \frac{n-r}{nm} = \frac{1}{m} - \frac{r}{nm} \to \frac{1}{m}\,.$$

Dunque,

$$\limsup \frac{a_n}{n} \le \frac{a_m}{m} \qquad \forall m\,.$$

E quindi

$$\limsup \frac{a_n}{n} \le \liminf \frac{a_n}{n}\,.$$

Ciò prova l'esistenza del limite. ∎

Esempio 193 Su $\mathbb{R}^2$ (riferito alla base canonica) consideriamo la trasformazione lineare descritta mediante la matrice

$$A = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}.$$

Per calcolare il raggio spettrale mediante la formula (191) bisogna prima di tutto calcolare le potenze di A:

$$A^{2n} = 2^n I\,, \qquad A^{2n+1} = 2^n A\,.$$

E' immediatamente evidente che $||A|| = 2$ e quindi

$$||A^{2n}||^{1/2n} = \sqrt{2}\,, \qquad ||A^{2n+1}||^{1/(2n+1)} = [\sqrt{2}]^{(2n+2)/(2n+1)}\,.$$

Dunque, il raggio spettrale è $\sqrt{2}$. Si noti che la successione $(||A^n||^{1/n})$ non è monotona. ∎

Osservazione 194 L'inverso che figura nella definizione di risolvente è in realtà l'inverso sinistro (si veda la Definizione 178). Al paragrafo 3.6.3.1 vedremo una classe importante di operatori per i quali l'immagine di $(zI - A)$ è chiusa quando $z \in \rho(A)$e quindi l'operatore $(zI - A)^{-1}$ è ovunque definito. ∎

2.12.1 Proiezioni spettrali

Sia X uno spazio di Banach e sia A un operatore, anche non continuo, da X in sé, con dominio denso.

Abbiamo notato che il risolvente è una funzione analitica e ciò suggerisce di studiare l'analogo, scritto per gli operatori, della formula integrale di Cauchy:

$$\frac{1}{2\pi i}\int_\gamma f(z)(zI-A)^{-1}\,\mathrm{d}z \tag{2.72}$$

ove γ è una curva semplice e chiusa[20] il cui sostegno è contenuto in $\rho(A)$. La funzione $f(z)$ è olomorfa.

Naturalmente, dovremo dare un senso all'integrale. Dato che la funzione a valori in $\mathcal{L}(X)$

$$z \to f(z)(zI-A)^{-1}$$

è uniformemente continua sul sostegno di γ, l'integrale si definisce coll'usuale metodo di Riemann, come "limite" delle somme di Riemann. Lasciamo al lettore i semplici dettagli.

Nonostante che la (2.72) abbia senso per ogni operatore lineare A da X in sé, purché il sostegno di γ sia contenuto in $\rho(A)$, noi ci limiteremo a considerare il cao degli operatori A continui.

Per interpretare la (2.72), consideriamo la funzione

$$f(z)=\sum_{n=0}^{+\infty} f_n z^n$$

e la serie corrispondente

$$\sum_{n=0}^{+\infty} f_n A^n\,. \tag{2.73}$$

Nel caso particolare in cui $f(z)=p(z)$ sia un polinomio, la serie (2.73) è una somma finita e definisce un operatore che, ovviamente, si indica col simbolo $p(A)$. Per esempio, se $p(z)=z^2$, allora $p(A)=A^2$. Se $f(z)$ è una generica funzione analitica la cui serie converge in un intorno di 0, la serie (2.73) in generale non converge, ma certamente converge in $\mathcal{L}(X)$ se

$$||A||<R$$

con R raggio di convergenza della serie di potenze di $f(z)$. Infatti in tal caso

$$\left|\left|\sum_{n=k}^{m} f_n A^n\right|\right| \le \sum_{n=k}^{m} f_n||A||^n \le \sum_{n=k}^{m} f_n r^n\,, \qquad r<R\,,$$

[20] come al solito, orientata in verso positivo, ossia antiorario.

e la convergenza si vede dal test di Cauchy per la convergenza delle serie.

Ricapitolando, se $||A|| < R$, la serie (2.73) converge e definisce un operatore di $\mathcal{L}(X)$, che indicheremo col simbolo

$$f(A)\,.$$

Ricordiamo ora che i coefficienti f_n si rappresentano come

$$f_n = \frac{1}{2\pi i}\int_\gamma \frac{f(\zeta)}{\zeta^{n+1}}\,\mathrm{d}\zeta$$

con γ curva semplice e chiusa orientata positivamente, il cui sostegno è contenuto nel disco di convergenza di $f(z)$.

Supponiamo che la curva γ racchiuda il disco $\{z \mid |z| < r\}$, con

$$||A|| < r < R\,. \tag{2.74}$$

In tal caso si trova

$$\begin{aligned} f(A) &= \sum_{n=0}^{+\infty} f_n A^n = \frac{1}{2\pi i}\sum_{n=0}^{+\infty}\int_\gamma f(\zeta)\frac{A^n}{\zeta^{n+1}}\,\mathrm{d}\zeta \\ &= \frac{1}{2\pi i}\int_\gamma f(\zeta)\left(\sum_{n=0}^{+\infty}\frac{A^n}{\zeta^{n+1}}\right)\mathrm{d}\zeta = \frac{1}{2\pi i}\int_\gamma f(\zeta)(\zeta I - A)^{-1}\,\mathrm{d}\zeta\,. \end{aligned}$$

Si noti che questo calcolo vale grazie alla condizione (2.74) e, se vale (2.74), allora si ha anche

$$\sigma(A) \subseteq \{z \mid \ |z| < R\}\,.$$

Osservazione 195 Si noti che l'integrale (2.72) ha senso anche se γ, di sostegno in $\rho(A)$, racchiude solo una parte dello spettro di A. Però in tal caso *non* useremo la notazione $f(A)$ per indicarlo. ∎

In un caso particolare è facile calcolare l'integrale (2.72): supponiamo che $\rho(A)$ contenga una regione di Jordan Ω e supponiamo che il sostegno di γ appartenga a Ω. In questo caso un argomento analogo a quello usato nella dimostrazione del teorema 189 (basato sul fatto che il duale distingue gli elementi dello spazio di Banach) prova che l'integrale è nullo. Ossia, il Teorema di Cauchy per le funzioni olomorfe vale anche per integrali della forma (2.72). Dunque, i casi interessanti saranno quelli nei quali γ "gira" intorno a punti di $\sigma(A)$. Per intuire cosa dobbiamo attenderci, consideriamo l'esempio seguente.

Ricordiamo che un operatore lineare P da uno spazio lineare X in se stesso si chiama PROIEZIONE quando $P^2 = P$. Nel caso di operatori tra spazi di Banach, sottintenderemo sempre che le "proiezioni" devono anche essere operatori continui.

Esempio 196 Sia $X = \mathbb{C}^3$ a sia

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad \text{così che} \qquad (zI - A)^{-1} = \begin{bmatrix} 1/z & 1/z^2 & 0 \\ 0 & 1/z & 0 \\ 0 & 0 & 1/(z-2) \end{bmatrix}.$$

Sia γ una curva semplice e chiusa che racchiude 0 e che lascia fuori 2. Si calcola immediatamente che

$$\frac{1}{2\pi i}\int_\gamma (zI - A)^{-1}\,\mathrm{d}z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Si trova così la proiezione sull'autospazio generalizzato dell'autovalore 0.

Operando in modo analogo con una curva che racchiude 2 e lascia fuori 0 si trova la proiezione sull'altro autospazio.

Abbiamo così calcolato l'integrale nel caso della funzione $f(z) = 1$. Se $f(z) = z$ un calcolo analogo dà

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

rispettivamente, a seconda della scelta della curva. Queste sono le restrizioni di A ai due autospazi. Si trova così una "diagonalizzazione a blocchi" della matrice A. ∎

Senza trattare l'integrale (2.72) in generale, vogliamo limitarci a considerare i due casi $f(z) = 1$ ed $f(z) = z$, che sono particolarmente importanti per le applicazioni, e che verranno usati nel paragrafo 3.8.4.

Generalizzando l'esempio 196, supponiamo che $\sigma(A) = \sigma_1(A) \cup \sigma_2(A)$ e che una regione di Jordan Ω contenga $\sigma_1(A)$ e lasci fuori $\sigma_2(A)$. Sia γ una curva semplice e chiusa col sostegno in Ω, che gira intorno a $\sigma_1(A)$, come nella figura 2.7, a sinistra:

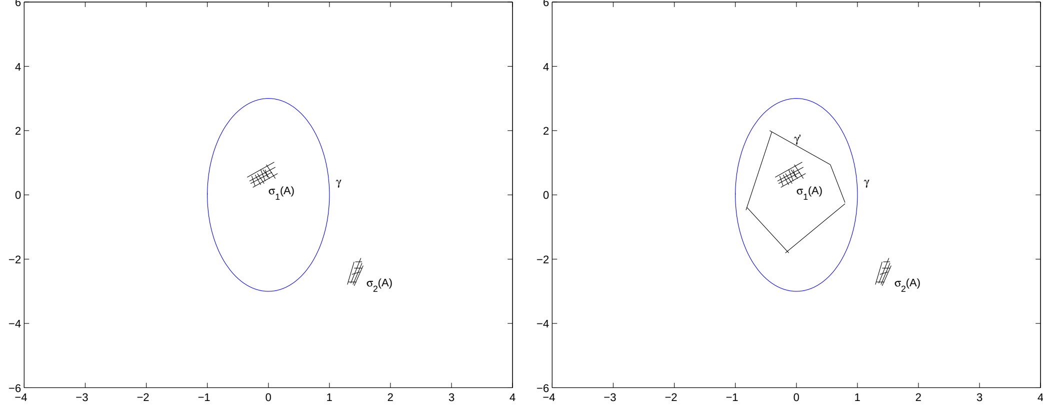
In tal caso:

Teorema 197 *Valgano le condizioni appena dette. L'operatore*

$$P = \frac{1}{2\pi i}\int_\gamma (zI - A)^{-1}\, dz$$

è una proiezione.

Figura 2.7:



<u>Dim</u>. Ricordiamo che vale il teorema di Cauchy. Da ciò si deduce che due curve γ e γ' semplici e chiuse in Ω, che ambedue racchiudono $\sigma_1(A)$ e lasciano fuori $\sigma_2(A)$ (come in figura 2.7, a destra) definiscono il medesimo operatore P:

$$P = \frac{1}{2\pi i}\int_\gamma (zI-A)^{-1}\,\mathrm{d}z = \frac{1}{2\pi i}\int_{\gamma'} (\zeta I-A)^{-1}\,\mathrm{d}\zeta\,.$$

Dunque,

$$\begin{aligned} P^2 &= \left\{\frac{1}{2\pi i}\int_\gamma (zI-A)^{-1}\,\mathrm{d}z\right\}\left\{\frac{1}{2\pi i}\int_{\gamma'} (\zeta I-A)^{-1}\,\mathrm{d}\zeta\right\} \\ &= \frac{1}{2\pi i}\cdot\frac{1}{2\pi i}\int_\gamma\left[\int_{\gamma'}(zI-A)^{-1}(\zeta I-A)^{-1}\right]\,\mathrm{d}\zeta\;\mathrm{d}z\,. \end{aligned}$$

Non è restrittivo supporre che la curva γ racchiuda la curva γ', come nella figura 2.7, a destra. A questo punto usiamo una formula[21] che si chiama PRIMA FORMULA DEL RISOLVENTE e che è di verifica immediata:

$$(zI-A)^{-1}(\zeta I-A)^{-1} = \frac{1}{\zeta - z}\left[(zI-A)^{-1} - (\zeta I-A)^{-1}\right]\,. \qquad (2.75)$$

[21]si noti che questa formula estende l'uguaglianza, valida tra numeri,

$$\frac{1}{(z-a)(\zeta-a)} = \frac{1}{\zeta-z}\left[\frac{1}{z-a} - \frac{1}{\zeta-a}\right]\,.$$

Usando questa formula, si trova

$$P^2 = \frac{1}{2\pi i}\int_\gamma \left[\frac{1}{2\pi i}\int_{\gamma'} \frac{1}{\zeta - z}(zI - A)^{-1}\,\mathrm{d}\zeta - \frac{1}{2\pi i}\int_{\gamma'} \frac{1}{\zeta - z}(\zeta I - A)^{-1}\,\mathrm{d}\zeta\right]\,\mathrm{d}z\,.$$

Ora, dal Teorema di Cauchy,

$$\frac{1}{2\pi i}\int_{\gamma'} \frac{1}{\zeta - z}(zI - A)^{-1}\,\mathrm{d}\zeta = 0$$

perché il punto z, che è sulla curva γ, è nella regione *esterna* a γ'.

Dunque rimane

$$\begin{aligned} P^2 &= -\frac{1}{2\pi i}\int_\gamma \left[\frac{1}{2\pi i}\int_{\gamma'} \frac{1}{\zeta - z}(\zeta I - A)^{-1}\,\mathrm{d}\zeta\right]\,\mathrm{d}z \\ &= \frac{1}{2\pi i}\int_{\gamma'} \left[\frac{-1}{2\pi i}\int_{\gamma} \frac{1}{\zeta - z}\,\mathrm{d}z\right](\zeta I - A)^{-1}\,\mathrm{d}\zeta\,. \end{aligned}$$

L'integrando

$$z \to \frac{1}{z - \zeta}$$

ha ζ per polo semplice, perché la curva γ' è racchiusa dalla curva γ. Dunque

$$\frac{-1}{2\pi i}\int_\gamma \frac{1}{\zeta - z}\,\mathrm{d}z = \frac{1}{2\pi i}\int_\gamma \frac{1}{z - \zeta}\,\mathrm{d}z = 1$$

e quindi

$$P^2 = \frac{1}{2\pi i}\int_{\gamma'} (\zeta I - A)^{-1}\,\mathrm{d}\zeta = P\,. \quad \blacksquare$$

Supponiamo ora che σ_1 e σ_2 sino due s.insiemi limitati di $\sigma(A)$, appartenenti alla regione interna rispettivamente di γ_1 e di γ_2, curve di Jordan di sostegno in $\rho(A)$ ed esterne l'una all'altra come in figura 2.8.

Poniamo

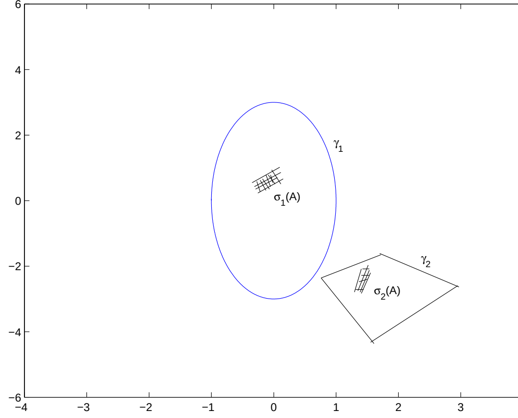
$$P_1 = \frac{1}{2\pi i}\int_{\gamma_1} (zI - A)^{-1}\,\mathrm{d}z\,, \qquad P_2 = \frac{1}{2\pi i}\int_{\gamma_2} (zI - A)^{-1}\,\mathrm{d}z\,.$$

Una dimostrazione analoga a quella del teorema precedente porta a:

Teorema 198 *Nelle ipotesi dette, si ha:* $P_1P_2 = P_2P_1 = 0$*. Inoltre* $X_1 = \operatorname{im} P_1$ *ed* $X_2 = \operatorname{im} P_2$ *sono spazi lineari invarianti per* A *e lo spettro della restrizione di* A *ad* $\operatorname{im} P_i$ *è l'insieme* σ_i*. Tale restrizione è data da*

$$Ax = \frac{1}{2\pi i}\int_{\gamma_i} z(zI - A)^{-1}x\,dz\,, \qquad \forall x \in X_i\,. \quad \blacksquare$$

Figura 2.8:



Omettiamo i dettagli della dimostrazione di questo teorema, che è analoga a quella del teorema 197.

Si noti che il Teorema 198 mostra una "diagonalizzazione a blocchi" dell'operatore A, analoga a quella vista nell'Esempio 196.

Nel corso della dimostrazione del Teorema 197 abbiamo introdotto la prima formula del risolvente (2.75) che si può interpretare come una decomposizione in "fratti semplici" della funzione

$$A \mapsto (zI - A)^{-1}(\zeta I - A)^{-1}\,.$$

D'altra parte questa formula studia il risolvente come funzione di λ per un fissato operatore A. Si può anche fissare λ e studiare il risolvente come funzione dell'operatore A. E' chiaro che ciò è più delicato perché operatori diversi possono avere insiemi risolventi privi di punti comuni. Quindi limitiamoci a considerare la funzione

$$A \mapsto (\lambda I - A)^{-1} \quad \text{definita su } \mathcal{L}(X)\,.$$

In tal caso lo spettro di A è contenuto nel disco $|\lambda| \leq \|A\|$ e quindi se A e B ambedue appartengono a $\mathcal{L}(X)$ i loro insiemi risolventi hanno intersezione non vuota.

Se A e B ambedue appartengono ad $\mathcal{L}(X)$ vale la SECONDA FORMULA DEL RISOLVENTE:

$$(\lambda I-A)^{-1}-(\lambda I-B)^{-1} = (\lambda I-A)^{-1}(A-B)(\lambda I-B)^{-1} \qquad \forall\lambda \in \sigma(A)\cap\sigma(B)\,.$$

Questa formula si vede immediatamente applicando $(\lambda I - A)^{-1}$ a sinistra e $(\lambda I - B)^{-1}$ a destra dei due membri dell'uguaglianza seguente:

$$(\lambda I - B) - (\lambda I - A) = A - B\,.$$

2.13 Trasformazioni non lineari

Fino ad ora abbiamo trattato soltanto di operatori lineari. Vogliamo ora presentare alcune considerazioni riguardanti i funzionali non lineari. Proviamo prima di tutto un teorema di punto fisso, ossia diamo una condizione per l'esistenza di soluzioni di un'equazione del tipo

$$f(x) = x\,.$$

In seguito, mostreremo come sia possibile estendere a funzioni tra spazi di Banach le definizioni di derivata e la prima formula degli incrementi finiti.

2.13.1 Teorema delle contrazioni e applicazioni

Supponiamo che f sia una trasformazione da uno spazio di Banach X in se stesso, non necessariamente lineare. Si dice che $f(x)$ è una CONTRAZIONE se esiste un numero $\alpha \in [0, 1)$ tale che

$$||f(x) - f(x')|| \le \alpha ||x - x'||\,.$$

Una contrazione è *lipschitziana* e quindi *continua.*

Se f è una qualsiasi trasformazione da X in sé, un punto $x_0 \in X$ si chiama PUNTO FISSO di f se

$$f(x) = x\,.$$

Vale:

Teorema 199 (delle CONTRAZIONI **o di**BANACH**)** *Sia K un insieme chiuso dello spazio di Banach X che è invariante per la contrazione $f(x)$. Esiste uno ed un solo punto fisso di $f(x)$ che appartiene a K.*

Dim. Proviamo prima di tutto che il punto fisso, se esiste, è *unico.* Siano per questo x ed y due punti fissi. Vale per essi

$$||x - y|| = ||f(x) - f(y)|| \le \alpha ||x - y||$$

ove α è *strettamente* minore di 1, per definizione di contrazione; e quindi la disuguaglianza precedente può solo valere se $x = y$.

Proviamo ora l'*esistenza* del punto fisso.

Fissiamo $k_0 \in K$ in modo arbitrario e costruiamo la successione

$$k_1 = f(k_0), \ldots, k_n = f(k_{n-1}).$$

Si noti che $k_n \in K$ per ogni n, perché $f(K) \subseteq K$.

Proveremo che (k_n) è una successione fondamentale e quindi convergente dato che lo spazio X è completo. Essendo K chiuso, il limite x_0 di (k_n) è in K. Passando al limite nei due membri dell'uguaglianza

$$k_n = f(k_{n-1})$$

si trova

$$x_0 = f(x_0)$$

e quindi x_0 è punto fisso di f.

Per completare la dimostrazione, basta mostrare che (k_n) è una successione fondamentale.

Stimiamo prima di tutto

$$||k_n - k_{n-1}|| = ||f(k_{n-1}) - f(k_{n-2})|| \le \alpha ||k_{n-1} - k_{n-2}||.$$

Iterando si vede che

$$||k_n - k_{n-1}|| \le \alpha^{n-1} ||k_1 - k_0||.$$

Valutiamo ora

$$\begin{aligned} &||k_{n+m} - k_n|| \le ||k_{n+m} - k_{n+m-1}|| + ||k_{n+m-1} - k_{n+m-2}|| + \cdots + ||k_{n+1} - k_n|| \\ &\le \{\alpha^{n+m-1} + \alpha^{n+m-2} + \cdots + \alpha^n\} ||k_1 - k_0|| \le \frac{\alpha^n}{1-\alpha} ||k_1 - k_0||. \end{aligned}$$

Ciò prova che la successione (k_n) è fondamentale e completa la dimostrazione. ∎

Osservazione 200 Sottolineiamo che la successione (k_n) costruita nella dimostrazione del teorema converge al punto fisso *per ogni* scelta del valore iniziale k_0. ∎

Presentiamo ora una semplice modifica del teorema 199 che è spesso utile. Indichiamo con $f^{(n)}$ la funzione su X ottenuta componendo f con sé stessa n–volte:

$$f^{(1)}(x) = f(x), \qquad f^{(k)}(x) = f\left(f^{(k-1)}(x)\right).$$

Può accadere che f non sia una contrazione, ma che esista un numero ν per cui $f^{(\nu)}$ è una contrazione. Vale:

Corollario 201 *Se f è continua e se $f^{(\nu)}$ è una contrazione su un s.insieme K chiuso di X tale che $f(K) \subseteq K$, allora $f(x)$ ammette un punto fisso in K e questo è unico.*

Dim. Notiamo che se $f(x_0) = x_0$ allora vale anche

$$f(f(x_0)) = f(x_0) = x_0$$

e quindi x_0 è anche punto fisso di $f^{(2)}$. Iterando questo procedimento, si vede che x_0 è anche punto fisso della contrazione $f^{(\nu)}$. Ciò mostra l'*unicità* del punto fisso. Proviamone ora l'*esistenza.*

Si sa che esiste il punto fisso x_0 di $f^{(\nu)}$:

$$x_0 = f^{(\nu)}(x_0)\,.$$

Applicando f ai due membri dell'uguaglianza si vede che

$$f(x_0) = f(f^{(\nu)}(x_0)) = f^{(\nu)}(f(x_0))$$

ossia, anche $f(x_0)$ è punto fisso della contrazione $f^{(\nu)}$. L'unicità del punto fisso implica che

$$f(x_0) = x_0\,. \;\blacksquare$$

Osservazione 202 Osserviamo che la ricerca dei punti fissi conduce anche alla ricerca di zeri di funzioni: il funto x_0 verifica $f(x_0) = 0$ se e solo se x_0 è punto fisso di $F(x) = x - f(x)$. ∎

2.13.1.1 Applicazioni: il metodo delle tangenti

E' noto il metodo delle tangenti per la determinazioni di zeri di funzioni convesse di variabile reale. Mostriamo come tale metodo si ritrovi mediante il teorema delle contrazioni. Sia per questo $f(x)$ convessa su $\mathbb{R}$ e di classe C^2. Supponiamo che la derivata prima non si annulli e supponiamo che sia

$$\left|\frac{f(x)f''(x)}{f'(x)^2}\right| \le \alpha < 1\,.$$

La funzione

$$F(x) = x - \frac{f(x)}{f'(x)}$$

ha un punto fisso x_0 se e solo se $f(x_0) = 0$ e viceversa (si ricordi che la derivata non si annulla).

Si calcola immediatamente che

$$F'(x) = \frac{f(x)f''(x)}{f'(x)^2}$$

e quindi, nelle ipotesi fatte, F è una contrazione. Ha quindi un punto fisso che si costruisce come segue: fissato un qualsiasi x_0, il punto x_1 è

$$x_1 = F(x_0) = x_0 - \frac{f(x_0)}{f'(x_0)}\,,$$

punto nel quale la tangente in $(x_0, f(x_0))$ al grafico di f,

$$y = f(x_0) + f'(x_0)(x - x_0)\,,$$

taglia l'asse delle ascisse.

Ciò è l'interpretazione geometrica del punto x_1 e quindi anche dei successivi punti x_n che approssimano lo zero di $f(x)$.

2.13.1.2 Equazioni integrali di Fredholm ed equazioni differenziali ordinarie

Sia $K(t, s, x)$ una funzione a valori reali continua su $[a, b] \times [a, b] \times \mathbb{R}$, lipschitziana nella terza variabile, uniformemente rispetto alla prima e alla seconda:

$$|K(t, s, x) - K(t, s, x')| \le M|x - x'|$$

con M indipendente da t e da s. Consideriamo la trasformazione T_μ da $C(a, b)$ in sé definita da

$$(T_\mu x)(t) = \mu \int_a^b K(t, s, x(s))\,\mathrm{d}s + f(t)$$

con $f(t)$ funzione continua fissata. E' chiaro che

$$|(T_\mu x)(t) - (T_\mu x')(t)| \le \mu \int_a^b M|x(s) - x'(s)|\,\mathrm{d}s \le \mu M(b - a)||x - x'||$$

e quindi la trasformazione T è una contrazione se

$$\mu M(b - a) < 1\,. \tag{2.76}$$

Dunque:

Teorema 203 *Se $\mu M(b-a) < 1$, l'equazione di Fredholm*

$$x(t) = \mu \int_a^b K(t, s, x(s))\, ds + f(t) \tag{2.77}$$

ammette soluzione e questa è unica.

Si noti che la condizione (2.76) puó realizzarsi o con $[a, b]$ fissato, prendendo μ piccolo, o con μ fissato, spesso $\mu = 1$, prendendo $b - a$ piccolo.

Le ipotesi di questo teorema possono indebolirsi e in particolare si vede che anche l'operatore

$$(T_\mu x)(t) = \mu \int_a^t K(t, s, x(s))\, \mathrm{d}s + f(t)$$

è una contrazione se $K(t, s, x)$ è continua per $a \leq s \leq t \leq b$ ed uniformemente lipschitziana in $x \in \mathbb{R}$. Nel caso particolare in cui $f(t) = x_0$, costante, e $\mu K(t, s, x) = K(s, x)$, l'equazione (2.77) equivale a

$$x'(t) = K(t, x(t))\,, \qquad x(a) = x_0\,. \tag{2.78}$$

Dunque,

Teorema 204 *Sia $K(t, x)$ continua in t, x ed uniformemente lipschitziana in x. Il problema di Cauchy (2.78) ammette soluzione su (a, b), con $|b - a|$ abbastanza piccolo, e tale soluzione è unica.*

2.13.2 I differenziali

Sia $f(x)$ una trasformazione da uno spazio di Banach X in uno spazio di Banach Y. Supponiamo che x_0 sia un punto interno al suo dominio.

Nel caso in cui $X = \mathbb{R}^n$ si sa che si possono definire le derivate direzionali e il differenziale in x_0. Vogliamo estendere queste definizioni al caso in cui X è un generico spazio di Banach.

Sia v un qualsiasi elemento di X. Consideriamo il limite

$$\lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t}\,.$$

Questo limite può esistere o meno. Se esiste si indica col simbolo

$$D_v f(x_0)$$

e si chiama la DERIVATA SECONDO IL VETTORE v. Se $||v|| = 1$ la derivata secondo il vettore v si chiama la DERIVATA DIREZIONALE nella direzione v.

La derivata direzionale può esistere in una direzione e non esistere in altre direzioni; e, se anche esiste in ogni direzione, la trasformazione

$$v \longrightarrow D_v f(x_0) \tag{2.79}$$

è in generale non lineare, come prova l'esempio seguente.

Esempio 205 Si definisce una funzione $f(x, y)$ su $\mathbb{R}^2$ come segue: prima di tutto si fissa una successione di punti (x_k, y_k) due a due distinti, tutti di norma 1, ossia tutti appartenenti alla circonferenza

$$C = \{(x, y) \mid x^2 + y^2 = 1\} .$$

Fissato un qualsiasi punto $(x, y) \neq (0, 0)$ di $\mathbb{R}^2$ si considera il punto

$$\frac{(x, y)}{||(x, y)||} .$$

Questo può essere uno dei punti (x_k, y_k) o meno. Se non è uno di tali punti, si pone $f(x, y) = 0$. Se invece esiste un indice k per cui

$$\frac{(x, y)}{||(x, y)||} = (x_k, y_k)$$

allora si definisce

$$f(x, y) = k||(x, y)|| = k\sqrt{x^2 + y^2} .$$

Inoltre definiamo $f(0, 0) = 0$.

Osserviamo che se $(x, y)/|(x, y)\| = (x_k, y_k)$ allora si ha anche $(tx, ty)/|(tx, ty)\| = (x_k, y_k)$ e quindi

$$f(tx, ty) = kt\sqrt{x^2 + y^2} . \tag{2.80}$$

Vogliamo calcolare $D_v f(0, 0)$

Fissata una qualsiasi direzione $v = (x, y)$, consideriamo i rapporti incrementali

$$\frac{f(tv)}{t} = \frac{f(tx, ty)}{t} .$$

Se $v/||v||$ non è uno dei punti (x_k, y_k), il valore del rapporto incrementale è zero per ogni t; e quindi

$$\lim_{t\to 0} \frac{f(tv)}{t} = 0 .$$

Altrimenti, se esiste un indice k_v per cui

$$\frac{v}{||v||} = k_v(x_{k_v}, y_{k_v})$$

allora, da (2.80),

$$\lim_{t\to 0} \frac{f(tv)}{t} = \lim_{t\to 0} \frac{tk_v||v||}{t} = k_v||v||\,.$$

Dunque, $df(x_0, v)$ esiste per ogni vettore v, ma non è funzione lineare di v. ∎

Quando invece l'operatore

$$v \longrightarrow D_v f(x_0)\,,$$

è *lineare,* questo si chiama il DIFFERENZIALE DI GÂTEAUX di f in x_0 e si indica col simbolo

$$df(x_0, v)\,.$$

Se esiste il differenziale di Gâteaux di f in x_0 allora, per ogni v fissato, si ha

$$\lim_{t\to 0} \left|\left| \frac{f(x_0 + tv) - f(x_0)}{t} - df(x_0, v) \right|\right| = 0 \tag{2.81}$$

e quindi

$$f(x_0 + tv) - f(x_0) = df(x_0, v)t + \mathrm{o}(t; x_0, v)$$

ove $\mathrm{o}(t; x_0, v)$ indica una funzione della variabile reale t a valori in X e tale che

$$\lim_{t\to 0} \frac{\mathrm{o}(t; x_0, v)}{t} = 0\,.$$

Si noti però che il limite in (2.81) *non* è generalmente uniforme rispetto a v. Si consideri infatti l'esempio seguente:

Esempio 206 Sia $X = \mathbb{R}^2$ e sia

$$f(x, y) = \begin{cases} 1 & \text{se } x^2 < x < x^4 \\ 0 & \text{altrimenti.} \end{cases}$$

Si vede facilmente che il differenziale di Gâteaux di questa funzione in $(0, 0)$ esiste e vale 0. Però, il limite (2.81) non è uniforme rispetto alla direzione. Infatti, sulla retta di pendenza m

$$x = t\,, \qquad y = mt$$

la disuguaglianza

$$\left|\frac{f(t,mt)}{t}\right| < \epsilon$$

vale quando $0 \leq t \leq \sqrt[3]{m}$. Dunque, non per $0 < t < \delta$ per un valore δ indipendente da m.

Si osservi che la funzione $f(x,y)$, pur essendo differenziabile secondo Gâteaux in $(0,0)$, non è continua. ∎

Si dice che una funzione $f(x)$ è DIFFERENZIABILE SECONDO FRÉCHET nel punto x_0 quando esiste un funzionale lineare continuo L per cui

$$f(x_0 + v) - f(x_0) = Lv + \mathrm{o}(v; x_0)\,.$$

Col simbolo $\mathrm{o}(v; x_0)$ si intende una funzione, questa volta da X in sé, tale che

$$\lim_{||v||\to 0} \frac{\mathrm{o}(v; x_0)}{||v||} = 0\,.$$

Si richiede cioè che L verifichi

$$\lim_{||v||\to 0} \frac{||f(x_0 + v) - f(x_0) - Lv||}{||v||} = 0\,.$$

Il funzionale lineare continuo L si chiama il *differenziale di Fréchet* della funzione f in x_0, e si indica col simbolo

$$df(x_0)v\,.$$

E' facile provare:

- Se esiste il differenziale di Fréchet in un punto x_0 allora esiste anche quello di Gâteaux, e questi coincidono;
- se esiste il differenziale di Fréchet nel punto x_0 allora la funzione è continua in x_0.

La formula

$$f(x_0 + v) - f(x_0) = df(x_0)v + \mathrm{o}(v; x_0)$$

generalizza la prima formula degli incrementi finiti.

Quando il differenziale di Fréchet esiste in ogni punto di un intorno di x_0, la funzione

$$x \to df(x)$$

si indica col simbolo $f'(x)$ e si chiama la FUNZIONE DERIVATA SECONDO FRÉCHET di $f(x)$. Questa funzione è generalmente non lineare ed opera da X a $\mathcal{L}(X,Y)$. Può ben essere che questa sia a sua volta differenziabile secondo Fréchet nei punti di un intorno di $x_0 \in X$. Si può quindi definirne la derivata prima e questa è, per definizione, la derivata seconda di f in x_0.

Procedendo analogamente, si definiscono anche le derivate successive.

Capitolo 3

Spazi di Hilbert

3.1 Prodotto interno e norma

Gli SPAZI DI HILBERT sono particolari spazi di Banach, che generalizzano $\mathbb{R}^n$ o $\mathbb{C}^n$ con l'usuale distanza euclidea.

Conviene introdurre prima di tutto la definizione di PRODOTTO INTERNO. Sia X uno spazio lineare. Si chiama prodotto interno su X una funzione $f(x,y)$ su $X \times X$, a valori nel campo scalare Φ (con $\Phi = \mathbb{R}$ oppure $\Phi = \mathbb{C}$), con queste proprietà:

- per ogni fissato y, la funzione $x \to f(x,y)$ è lineare:

$$f(\alpha x + \beta x', y) = \alpha f(x,y) + \beta f(x', y)\,.$$

- per ogni x ed y vale $f(x,y) = \overline{f(y,x)}$. Questa proprietà implica in particolare che la parte immaginaria di $f(x,x)$ è nulla per ogni x.

- vale $f(x,x) > 0$ per ogni $x \neq 0$.

La prima proprietà mostra che

$$f(0,0) = f(r \cdot 0, 0) = r f(0,0)$$

per ogni numero r; e quindi, scegliendo $r = 0$,

$$f(0,0) = 0\,.$$

La terza proprietà richiede che $f(x,x) = 0$ valga *solo* per $x = 0$.

Si noti che la funzione $f(x,y)$ *non* è lineare rispetto ad y ma, per ogni fissato x, vale

$$\begin{aligned} & f(x,\alpha y+\beta y')=\overline{f(\alpha y+\beta y',x)}=\overline{\alpha f(y,x)+\beta f(y',x)} \\ & =\bar\alpha f(x,y)+\bar\beta f(x,y')\,. \end{aligned} \tag{3.1}$$

Le proprietà (3.1) si chiama ANTILINEARITÀ.

Se accade che $\Phi=\mathbb{R}$ allora gli scalari sono reali e quindi si ha linearità anche nella seconda componente.

In pratica per indicare il prodotto interno di x ed y si usa il simbolo $\langle x,y\rangle$ (o simboli analoghi, per esempio $\langle x|y\rangle$, $(x|y)$ o anche (x,y)). Noi useremo il simbolo $\langle x,y\rangle$, notando la somiglianza col simbolo $\langle\langle x^*,x\rangle\rangle$ usato per rappresentare l'azione del funzionale lineare x^* su x. Si noti però che $\langle\langle x^*,x\rangle\rangle$ è lineare sia rispetto alla prima che alla seconda variabile, anche quando $\Phi=\mathbb{C}$.

Due vettori x ed y si dicono ORTOGONALI quando il loro prodotto interno è nullo:

$$x\perp y \iff \langle x,y\rangle=0\,.$$

Proviamo che per i prodotti interni vale la DISUGUAGLIANZA DI SCHWARZ:

Teorema 207 *Per ogni x,y vale*

$$|\langle x,y\rangle|^2\le\langle x,x\rangle\langle y,y\rangle\,.$$

L'uguaglianza vale se e solo se i vettori x ed y sono colineari, ossia se e solo se $x=\alpha y$, $\alpha\in\Phi$.

Dim. Se $\langle x,y\rangle=0$ allora la disuguaglianza è ovvia. Supponiamo quindi esplicitamente $\langle x,y\rangle\neq0$ e introduciamo

$$a=\frac{\langle x,y\rangle}{|\langle x,y\rangle|}\,.$$

Consideriamo quindi che per ogni t (reale o complesso) vale

$$0\le\langle\bar a x+ty,\bar a x+ty\rangle\,.$$

Scegliamo t reale e introduciamo in quest'espressione la definizione di a. Si trova

$$0\le\langle\bar a x+ty,\bar a x+ty\rangle=\langle y,y\rangle t^2+2|\langle x,y\rangle|t+\langle x,x\rangle\,. \tag{3.2}$$

Questo è un polinomio in t, a coefficienti reali. Il segno di questo polinomio è costante e quindi il suo discriminante è negativo, ossia:

$$|\langle x,y\rangle|^2\le\langle x,x\rangle\langle y,y\rangle\,. \tag{3.3}$$

Questa è la disuguaglianza che volevamo provare.

Se in (3.3) vale l'uguaglianza, allora il polinomio (3.2) è un quadrato:

$$\langle \bar{a}x + ty, \bar{a}x + ty \rangle = (mt + n)^2$$

per certi numeri m ed n. E' quindi nullo per $t = -n/m$, ossia

$$\langle \bar{a}x + (-n/m)y, \bar{a}x + (-n/m)y \rangle = 0 \qquad \text{e dunque} \qquad \bar{a}x + (-n/m)y = 0\,,$$

così che i vettori x ed y sono colineari. ∎

Teorema 208 *La funzione definita su X da*

$$x \to \sqrt{\langle x, x \rangle}$$

è una norma su X.

<u>Dim</u>. Usando la disuguaglianza di Schwarz, proviamo che vale la disuguaglianza triangolare:

$$\begin{aligned}
&\langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\
&= \langle x, x \rangle + 2\Re e \, \langle x, y \rangle + \langle y, y \rangle \\
&\le \langle x, x \rangle + 2|\langle x, y \rangle| + \langle y, y \rangle \qquad \text{(usando la disuguaglianza di Schwarz)} \\
&\le \langle x, x \rangle + 2\left(\langle x, x \rangle\right)^{1/2}\left(\langle y, y \rangle\right)^{1/2} + \langle y, y \rangle = \left[\langle x, x \rangle^{1/2} + \langle y, y \rangle^{1/2}\right]^2 .
\end{aligned}$$

Si ha quindi

$$\sqrt{\langle x + y, x + y \rangle} \le \sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle}\,.$$

Questa è la disuguaglianza triangolare. Le altre proprietà della norma sono immediate. Si noti che la proprietà $||x|| > 0$ per $x \neq 0$ vale perché $\langle x, x \rangle \neq 0$ per $x \neq 0$. ∎

Naturalmente scriveremo

$$||x|| = \sqrt{\langle x, x \rangle}\,. \tag{3.4}$$

Con questa notazione, la disuguaglianza di Schwarz si scrive

$$|\langle x, y \rangle \le ||x|| \cdot ||y||\,.$$

E' conseguenza della disuguaglianza di Schwarz e della definizione di norma l'asserto seguente:

Corollario 209 *Per ogni $y \in X$ fissato, il funzionale lineare*

$$x \to \langle x, y \rangle$$

è continuo sullo s.l.n. X, dotato della norma (3.4)

Dim. Infatti, dalla disuguaglianza di Schwarz,

$$|\langle x, y \rangle| \leq M ||x|| \,, \qquad \text{con} \quad M = ||y|| \,. \;\blacksquare$$

Le norme che discendono da un prodotto interno godono di una proprietà bene particolare:

Teorema 210 *Sia $||x|| = \sqrt{\langle x, x \rangle}$. Questa particolare norma verifica l'uguaglianza*

$$||x + y||^2 + ||x - y||^2 = 2 \left(||x||^2 + ||y||^2 \right) \tag{3.5}$$

Dim. Si calcola immediatamente

$$\begin{aligned} &||x + y||^2 + ||x - y||^2 = \langle x + y, x + y \rangle + \langle x - y, x - y \rangle \\ &= ||x||^2 + \langle x, y \rangle + \langle y, x \rangle + ||y||^2 + ||x||^2 - \langle x, y \rangle - \langle y, x \rangle + ||y||^2 \\ &= 2 \left(||x||^2 + ||y||^2 \right) \,. \;\blacksquare \end{aligned}$$

L'uguaglianza (3.5) si chiama IDENTITÀ DEL PARALLELOGRAMMA. Nella geometria piana essa si enuncia dicendo che *la somma dei quadrati costruiti sulle diagonali di un parallelogramma è uguale alla somma dei quadrati costruiti sui lati.*

E' importante sapere che non tutte le norme discendono da un prodotto interno. Infatti vale:

Esempio 211 Si doti $\mathbb{R}^2$ della norma

$$||(\xi, \eta)|| = \max\{|\xi| \,,\; |\eta|\} \,.$$

Si provi che l'identità del parallelogramma non vale per la coppia dei vettori $x = (1, 0)$ ed $y = (0, 1)$. $\blacksquare$

Quest'osservazione suggerisce di dare un nome particolare agli s.l.n-ti la cui norma proviene da un prodotto interno. Questi si chiamano SPAZI PREHILBERTIANI e, se sono anche *completi,* si chiamano SPAZI DI HILBERT.

Lavoreremo ora esclusivamente con spazi di Hilbert, che indicheremo genericamente col simbolo H. I suoi elementi si indicano con i simboli h, k . . . .

E' chiaro che ogni spazio di Hilbert è anche uno speciale spazio di Banach. La proprietà di essere spazio di Hilbert è particolarmente importante per le applicazioni. Per esempio, possiamo notare subito che *in spazi prehilbertiani vale il* TEOREMA DI PITAGORA*:*

Teorema 212 *Siano h, k due elementi tra loro ortogonali di uno spazio di prehilbertiano H. Vale:*

$$||h+k||^2 = ||h||^2 + ||k||^2 .$$

Dim. Si calcola immediatamente

$$||h+k||^2 = \langle h+k, h+k\rangle = ||h||^2 + \langle h,k\rangle + \langle k,h\rangle + ||k||^2 = ||h||^2 + ||k||^2$$

perché $h \perp k$. ∎

3.1.1 Esempi di prodotti interni e di spazi di Hilbert

Elenchiamo gli spazi di Hilbert di uso più comune. Naturalmente essi si sono già incontrati come particolari spazi di Banach.

Il paragrafo 2.7 riporta un elenco di spazi di Banach e ne rappresenta in modo “concreto” i duali. Può essere opportuno, ma non è indispensabile, conoscere queste rappresentazioni per la lettura di questo paragrafo.

3.1.1.1 Lo spazio euclideo ad n dimensioni

è uno spazio di Hilbert, con prodotto interno

$$\langle h,k\rangle = \sum_{i=1}^{n} \bar{k}_i h_i \qquad \text{se} \quad h = \operatorname{col}[h_i]\,, \quad k = \operatorname{col}[k_i]\,.$$

3.1.1.2 Lo spazio l^2

è uno spazio di Hilbert, dotato del prodotto interno

$$\langle (h_i),(k_i)\rangle = \sum_{i=i}^{+\infty} \bar{k}_i h_i \,.$$

La convergenza della serie, quando h_i e k_i sono in l^2, è stata provata nel paragrafo 2.7.0.1. Possiamo ora notare che la convergenza segue applicando la disuguaglianza di Schwarz alle somme finite, e passando al limite.

Si ricordi che il duale di l^2 è isometrico a l^2 stesso.

3.1.1.3 lo spazio $L^2(K)$

è uno spazio di Hilbert, il cui prodotto interno è

$$\langle f, g\rangle = \int_K \bar{g}(x) f(x)\, \mathrm{d}x\,.$$

L'integrale dipende dagli elementi di $L^2(K)$, ossia dalle classi di equivalenza, e non dai rappresentanti delle classi stesse, e converge grazie alla disuguaglianza di Schwarz per gli integrali.

Si ricordi, dal paragrafo 2.7.0.3, che anche in questo caso lo spazio è una realizzazione del suo duale.

3.1.1.4 Lo spazio H^2

è uno spazio di Hilbert. Il prodotto interno nel caso di $H^2(D)$ è[1]

$$\langle f, g\rangle = \sup_{r\in(0,1)} \left[\int_0^{2\pi} \bar{f}(re^{it}) g(re^{it})\, \mathrm{d}t\right].$$

Nel caso di $H^2(\Pi)$ il prodotto interno è

$$\langle f, g\rangle = \sup_{x>0} \left[\int_{-\infty}^{+\infty} \bar{f}(x+iy) g(x+iy)\, \mathrm{d}y\right].$$

3.1.1.5 Lo spazio $W^{12}(K)$

è uno spazio di Hilbert dotato del prodotto interno

$$\langle f, g\rangle = \int_K \bar{g}(x) f(x)\, \mathrm{d}x + \int_K \nabla \bar{f}(x) \cdot \nabla g(x)\, \mathrm{d}x\,.$$

Nel caso in cui $K = [a, b]$, un prodotto interno che conduce ad una norma equivalente è

$$\langle f, g\rangle = \bar{g}(a) f(a) + \int_a^b \bar{g}'(x) f'(x)\, \mathrm{d}x\,.$$

[1] ricordiamo che frequentemente la norma si definisce premettendo un fattore $1/2\pi$ all'integrale.

3.1.1.6 Spazi L^2 ed l^2 "con peso"

Sia $\rho(x)$ una funzione misurabile non negativa definita su un insieme misurabile Ω. Indichiamo con $\mathcal{L}^2(K;\rho)$ l'insieme delle funzioni misurabili $f(x)$ tali che

$$\int_\Omega \rho(x)|f(x)|^2 \,\mathrm{d}x < +\infty \,.$$

La disuguaglianza di Minkowski mostra che questo è uno spazio lineare. Introducendo la relazione di equivalenza

$$f \sim g \qquad \text{se} \qquad \int_\Omega \rho(x)|f(x) - g(x)|^2 \,\mathrm{d}x = 0$$

si vede facilmente che lo spazio quoziente può essere dotato della struttura di spazio di Hilbert, esattamente come si fa quando si introduce lo spazio $L^2(\Omega)$ che, con queste notazioni, è niente altro che $L^2(\Omega;\rho)$ con $\rho(x) \equiv 1$.[2]

Il prodotto interno in $L^2(\Omega,\rho)$ è

$$\langle f, g\rangle = \int_\Omega \rho(x)\bar{g}(x)f(x)\,\mathrm{d}x$$

a cui corrisponde la norma

$$||f||^2 = \int_\Omega \rho(x)|f(x)|^2 \,\mathrm{d}x$$

(si noti che $\rho(x)$ non è elevata al quadrato).

In questo contesto la funzione $\rho(x)$ si chiama FUNZIONE PESO e lo spazio $L^2(\Omega,\rho)$ si chiama spazio "L^2 con peso ρ".

Si noti che la funzione $\rho(x)$ potrebbe tendere a zero, o addirittura essere identicamente zero su un s.insieme di Ω, e potrebbe essere illimitata.

Si vede facilmente che se esistono costanti m ed M per cui

$$0 < m \leq \rho(x) \leq M$$

allora lo spazio $L^2(\Omega;\rho)$ ha gli stessi elementi di $L^2(\Omega)$, e le norme di $L^2(\Omega;\rho)$ e di $L^2(\Omega)$ si equivalgono.

In modo analogo si può definire lo spazio l^2 "con peso ρ". In questo caso $\rho = (\rho_n)$ è una successione non negativa e gli elementi di tale spazio sono ora le successioni (x_n) tali che

$$||(x_n)||^2 = \sum_{n=0}^{+\infty} \rho_n |x_n|^2 < +\infty \,.$$

[2]Naturalmente, procedendo in modo analogo, si potrebbero introdurre gli spazi di Banach $L^p(\Omega,\rho)$ per ogni $p \geq 1$.

3.1.2 Uno spazio di Hilbert non separabile

Tutti gli esempi precedenti sono esempi di spazi di Hilbert separabili. Mostriamo un esempio di spazio di Hilbert *non* separabile. Osserviamo che se $||x|| = ||y|| = 1$ e se $x \perp y$, allora

$$||x - y||^2 = 2\,,$$

ossia x dista $\sqrt{2}$ da y. Dunque, *se in uno spazio di Hilbert si trova una famiglia non numerabile di vettori a due a due ortogonali, questo spazio non è separabile.*

Per ogni s *reale* consideriamo la funzione

$$\chi_s(t) = e^{ist}\,, \qquad t \in \mathbb{R}\,.$$

Consideriamo lo spazio lineare $\mathcal{S}$ generato da queste funzioni e su esso il prodotto interno

$$\langle f, g\rangle = \lim_{T\to+\infty} \frac{1}{2T}\int_{-T}^{T} \bar{g}(t) f(t)\,\mathrm{d}t\,.$$

E' facile vedere che si è definito un prodotto interno e quindi una norma su $\mathcal{S}$ e che $\|\chi_s\| = 1$ per ogni s.

Dunque $\mathcal{S}$ uno spazio prehilbertiano. Il suo completamento, introdotto nel teorema 29, è quindi uno spazio di Hilbert che non è separabile perchè se $s \neq r$

$$\langle \chi_s, \chi_r\rangle = \lim_{T\to+\infty} \frac{1}{2T}\int_{-T}^{T} e^{i(s-r)t}\,\mathrm{d}t = \lim_{T\to+\infty} \frac{e^{i(s-r)T} - e^{-i(s-r)T}}{2T(s-r)} = 0\,.$$

Dunque in questo spazio c'è un sistema *non numerabile* di vettori due a due ortogonali e di norma 1. Come si è detto, ciò basta a mostrare che lo spazio *non è separabile.*

3.2 Teorema delle proiezioni

Gli spazi di Hilbert, come si è notato, sono particolari spazi di Banach, dotati di proprietà speciali, utili per le applicazioni. Essenzialmente esse discendono tutte dal TEOREMA DELLE PROIEZIONI, che è in realtà un complesso di affermazioni che è bene studiare separatamente. In particolare è bene essere precisi, distinguendo le affermazioni che valgono in spazi prehilbertiani da quelle che richiedono la completezza.

Sia H uno spazio prehilbertiano e sia X un suo s.spazio. Sia $h \in H$. Un punto $x_0 \in X$ si chiama PROIEZIONE ORTOGONALE DI h SU X se

$$h - x_0 \perp x \qquad \forall x \in X\,.$$

Per indicare che $h - x_0$ è perpendicolare *ad ogni* elemento di X, scriveremo anche

$$h - x_0 \perp X\,.$$

Si noti che se $h \in X$ allora h è proiezione di se stesso su X, $h = x_0$.

In un generico spazio di Banach, una definizione analoga non può darsi perché l'ortogonalità non è definita. Anche in spazi di Hilbert però non è affatto ovvio che, dato h, la sua proiezione x_0 su X debba esistere. Se però essa esiste allora si può scrivere

$$h = (h - x_0) + x_0$$

e $h-x_0$, essendo perpendicolare ad X, è in particolare perpendicolare a $x_0 \in X$. Dunque, usando il *teorema di Pitagora,* si ha

Teorema 213 *Sia $h \in H$, H uno spazio prehilbertiano, ed esista la proiezione x_0 di h su X. Vale:*

$$||h||^2 = ||h - x_0||^2 + ||x_0||^2\,.$$

In particolare,

$$||x_0|| \le ||h||\,, \qquad ||h - x_0|| \le ||h||\,.$$

Abbiamo detto che l'esistenza della proiezione non è ovvia. Possiamo però immediatamente provare che, se la proiezione esiste, essa è unica:

Teorema 214 *Sia h un elemento dello spazio prehilbertiano H. Sia X un s.spazio di H. Se esiste, la proiezione di h su X è unica.*

Dim. Siano infatti x_0 ed x_1 due proiezioni di h su X. In tal caso,

$$\langle h - x_0, x\rangle = 0\,, \qquad \langle h - x_1, x\rangle = 0 \qquad \forall x \in X\,.$$

Usando la linearità della prima componente del prodotto interno si trova

$$\langle x_1 - x_0, x\rangle = 0 \qquad \forall x \in X\,.$$

Ora, X è uno spazio lineare a cui appartengono sia x_0 che x_1 e quindi anche $x_1 - x_0 \in X$. Scegliendo $x = x_1 - x_0$ si trova

$$0 = \langle x_1 - x_0, x_1 - x_0\rangle = ||x_1 - x_0||^2$$

e quindi $x_1 = x_0$. ∎

Il problema della proiezione è uno dei problemi che si studiano nella geometria euclidea e si sa che, in tale contesto, la proiezione x_0 di h è anche il punto di X che ha minima distanza da h. Questa proprietà vale anche in spazi prehilbertiani:

Teorema 215 *Sia H uno spazio prehilbertiano e sia X un suo sottospazio. Un punto $x_0 \in X$ è proiezione su X di $h \in H$ se e solo se*

$$||h - x_0|| \le ||h - x|| \qquad \forall x \in X\,.$$

Dim. Sia x_0 la proiezione di h su X e sia $x \in X$ qualsiasi. Si scriva

$$h - x = (h - x_0) + (x_0 - x)\,.$$

Essendo $(h - x_0) \perp (x - x_0)$, dal *teorema di Pitagora* se gue

$$||h - x||^2 = ||h - x_0||^2 + ||x - x_0||^2 \ge ||h - x_0||^2\,.$$

Ciò prova che x_0 è punto di minima distanza.

Viceversa, sia

$$||h - x_0|| \le ||h - x|| \qquad \forall x \in X\,.$$

Mostriamo che x_0 è proiezione di h su X, ossia che

$$\langle h - x_0, x\rangle = 0 \qquad \forall x \in X\,. \tag{3.6}$$

Si fissi un qualsiasi $x \in X$ e si consideri la funzione

$$t \to ||(h - x_0 + tx||^2 = ||h - x_0||^2 + t(2\Re e\,\langle h - x_0, x\rangle) + t^2||x||^2\,.$$

Questa è un polinomio nella variabile reale t, che ha minimo per $t = 0$. Se il campo scalare è $\mathbb{R}$, uguagliando a zero la derivata prima calcolata per $t = 0$ si trova la (3.6). Se il campo scalare è $\mathbb{C}$, si trova

$$\Re e\,\langle h - x_0, x\rangle = 0\,.$$

Si noti però che $x \in X$ è qualsiasi e quindi la precedente vale anche sostituendo x con ix. Così facendo si trova che vale anche

$$\mathcal{I}m\,\langle h - x_0, x\rangle = 0\,.$$

La (3.6) segue combinando queste due. ∎

Osservazione 216 Si osservi che la proprietà di minima distanza può anche introdursi in un generico spazio di Banach. Però in generale il punto di X che meno dista da h, se H *non* è uno spazio di Hilbert, né esiste né è unico.

Rinunciamo a presentare un esempio che mostra la non esistenza e mostriamo la non unicità. Sia per questo H lo spazio $\mathbb{R}^2$, ma dotato della norma

$$||x|| = ||(\xi, \eta)|| = \max\{|\xi|\,,\ |\eta|\}\,.$$

Sia $X = \{(\xi, 0) \mid \xi \in \mathbb{R}\}$ l'asse delle ascisse e sia $h = (0, 1)$. Si vede facilmente che

$$||h - x|| = 1 \qquad \forall x = (\xi, 0)\,, \quad \xi \in [-1, 1]\,.$$

Se invece $x = (\xi, 0)$ con $|\xi| > 1$ allora

$$||h - x|| = |\xi| > 1\,.$$

Dunque, il punto dell'asse delle ascisse che ha minima distanza da h non è unico e i punti di minima distanza sono quelli del segmento $[-1, 1]$. ∎

Esaminiamo ora il problema di minimo

$$\min\{||h - x|| \mid x \in X\}\,.$$

In generale, un problema di minimo non ha soluzione, ma si possono sempre costruire "successioni minimizzanti". Nel caso nostro, sia

$$d = \inf\{||h - x||\ x \in X\}$$

e, per ogni n, sia x_n tale che

$$d \le ||h - x_n|| < d + 1/n\,. \tag{3.7}$$

Proviamo:

Teorema 217 *Sia H uno spazio prehilbertiano. La successione (x_n) è fondamentale.*

Dim. Fissati n ed m, si deve valutare $||x_n - x_m||$. Per semplicità valutiamone il quadrato. *Usiamo l'identità del parallelogramma* per scrivere

$$\begin{aligned} &||x_n - x_m||^2 = ||(x_n - h) + (h - x_m)||^2 \\ &= 2\left[||x_n - h||^2 + ||h - x_m||^2\right] - ||(x_n - h) - (h - x_m)||^2 \\ &= 2\left[||x_n - h||^2 + ||h - x_m||^2\right] - 4||\frac{x_n + x_m}{2} - h||^2\,. \end{aligned}$$

E'

$$\frac{1}{2}(x_n + x_m) \in X$$

e quindi

$$||\frac{x_n + x_m}{2} - h||^2 > d^2\,.$$

Dalla definizione di (x_n), assegnato $\epsilon > 0$, segue l'esistenza di N_ϵ tale che, se n, m sono maggiori di N_ϵ, si ha

$$||x_n - h||^2 < d^2 + \epsilon/4\,, \qquad ||x_m - h||^2 < d^2 + \epsilon/4\,.$$

Dunque, per n, m maggiori di N_ϵ vale anche

$$||x_n - x_m||^2 < 2\left[2d^2 + \frac{\epsilon}{2}\right] - 4||\frac{x_n + x_m}{2} - h||^2 \leq 4d^2 + \epsilon - 4d^2 = \epsilon\,.$$

La successione (x_n) è quindi fondamentale. ∎

Di conseguenza:

Teorema 218 *Sia H uno* spazio di Hilbert *e sia X un suo s.spazio* chiuso. *Per ogni $h \in H$ esiste x_0, proiezione di h su X.*

Dim. Si costruisce la successione (x_n), definita da (3.7). Si sa che questa è una successione fondamentale in H, e quindi convergente, perché H è *completo.*

Sia

$$x_0 = \lim x_n\,.$$

Per ogni n, si ha $x_n \in X$ e quindi $x_0 \in X$ perché X è *chiuso.*

Da (3.7) si ha

$$d = \lim ||h - x_n||\,.$$

D'altra parte la *continuità della norma* mostra che

$$||h - x_0|| = \lim ||h - x_n||$$

e quindi x_0 è punto di minima distanza; e quindi è la proiezione di h su X. ∎

3.3 Insieme ortogonale, proiezioni ortogonali e complementi ortogonali

Sia A un qualsiasi s.insieme non vuoto di uno spazio di Hilbert H. Definiamo

$$A^{\perp} = \{h \mid h \perp A\} = \{h \mid \langle h, a\rangle = 0\ \forall a \in A\}\,.$$

Questa definizione dice che ogni elemento di $A^{\perp}$ è ortogonale ad ogni elemento di A e quindi che si ha anche

$$A \subseteq \left[A^{\perp}\right]^{\perp}\,. \tag{3.8}$$

In generale, diciamo che due s.insiemi A e B non vuoti di H sono tra loro ortogonali quando

$$\langle a, b\rangle = 0 \qquad \forall a \in A\,,\ \forall b \in B\,.$$

In tal caso scriviamo

$$A \perp B\,.$$

Dunque in particolare $A \perp A^{\perp}$.

Ovviamente:

Lemma 219 *Per ogni insieme A vale*

$$A \cap A^{\perp} = \{0\}\,.$$

<u>Dim</u>. Se infatti $a \in A \cap A^{\perp}$ allora $\langle a, a\rangle = 0$ e quindi $a = 0$. ∎

Vale:

Lemma 220 *Siano A e B due s.insiemi di H, con $A \subseteq B$. Allora $A^{\perp} \supseteq B^{\perp}$.*

<u>Dim</u>. Per definizione, $h \perp B$ quando $\langle h, b\rangle = 0$ per ogni $b \in B$. Essendo $A \subseteq B$, se $h \in B^{\perp}$ si ha $\langle a, h\rangle = 0$ e quindi

$$h \in B^{\perp} \implies h \in A^{\perp} \quad \text{ossia } B^{\perp} \subseteq A^{\perp}\,. \ ∎$$

Inoltre:

Teorema 221 *L'insieme $A^{\perp}$ è un s.spazio chiuso di H. Se A è denso in H allora $A^{\perp} = \{0\}$.*

Se $A^{\perp} = \{0\}$ e se A è un s.spazio, allora A è denso in H.

<u>**Dim.**</u> Siano x, y elementi di $A^\perp$ e siano α e β scalari. Per ogni $a \in A$ vale

$$\langle \alpha x + \beta y, a\rangle = \alpha\langle x, a\rangle + \beta\langle y, \alpha\rangle = 0\,.$$

Ciò prova che $A^\perp$ è un s.spazio (anche se A non lo è.)

Per provare che $A^\perp$ è chiuso, sia (x_n) una successione di elementi di $A^\perp$ e supponiamo che essa converga ad x_0. Dobbiamo provare che $x_0 \in A^\perp$. La continuità del prodotto interno mostra che, per ogni $a \in A$,

$$\langle x_0, a\rangle = \lim\langle x_n, a\rangle = 0\,.$$

Dunque, $x_0 \in A^\perp$, come volevamo.

Sia ora A denso in H e sia $x \in A^\perp$. Mostriamo che

$$\langle x, h\rangle = 0 \tag{3.9}$$

per ogni $h \in H$. Da ciò, scegliendo in particolare $h = x$, seguirà $x = 0$. Proviamo quindi che vale (3.9). Se accade che $h \in A$, allora vale (3.9). Se $h \notin A$, essendo A denso, esiste una successione (a_n) in A, convergente ad h. Dunque, ancora per la continuità del prodotto interno,

$$\langle x, h\rangle = \lim\langle x, a_n\rangle = 0\,.$$

Ricapitolando, abbiamo provato che se A è denso in H allora $A^\perp = \{0\}$.

Viceversa sia $A^\perp = \{0\}$ e sia inoltre A un *s.spazio*. Mostriamo che A è denso in H. Procedendo per assurdo, se il s.spazio A non è denso in H, la sua chiusura X è un s.spazio chiuso che non contiene un elemento $h \in H$. Sia x_0 la proiezione di h su X. Il vettore $h - x_0$ è non nullo, ed ortogonale ad X e quindi anche ad A. ∎

Osservazione 222 La condizione $A^\perp = \{0\}$ da sola, ossia senza l'ipotesi che A sia anche un s.spazio, *non implica che A debba essere denso in H.* Infatti, se $A = \{x \,:\, \|x\| = 1\}$ allora A non è denso in H ma si vede facilmente che $A^\perp = \{0\}$ perchè se $x \perp A$ ed $x \neq 0$ si hanno le due uguaglianze contrastanti

$$\|x\| \neq 0\,, \qquad 0 = \langle x, x/\|x\|\rangle = \|x\|\,. \ ∎$$

Ricordiamo ora la definizione

$$\operatorname{span} A = \left\{ \sum_{\text{finita}} \alpha_i a_i\,, \quad a_i \in A\,,\ \alpha_i \in \mathbb{C} \right\}.$$

Lo spazio lineare span A generalmente non è chiuso. Notiamo che:

Lemma 223

$$A^{\perp} = (\mathrm{cl}\,\mathrm{span}\,A)^{\perp}\ .$$

Dim. L'inclusione $(\mathrm{span}\,A)^{\perp} \subseteq A^{\perp}$ vale perché $A \subseteq \mathrm{span}\,A$. Inoltre,

$$(\mathrm{cl}\,\mathrm{span}\,A)^{\perp} = (\mathrm{span}\,A)^{\perp}$$

per la continuità del prodotto interno.

L'inclusione opposta vale perché se $h \in A^{\perp}$ allora $h \in (\mathrm{span}\,A)^{\perp}$ per la linearità del prodotto interno e quindi vale anche $h \in \mathrm{cl}\,[\mathrm{span}\,A]$ per la continuità del prodotto interno. ∎

Quest'ultima osservazione mostra che studiando gli insiemi perpendicolari ci si può limitare a studiare i s.insiemi di H che sono s.spazi chiusi. A questo riguardo notiamo:

Teorema 224 *Se X è un s.spazio chiuso di H allora*

$$H = X \oplus X^{\perp}\,.$$

Dim. Infatti, sia $h \in H$ e sia x la sua proiezione ortogonale su X. Allora,

$$x \in X\,, \quad h - x \in X^{\perp}\,, \qquad h = x + (h - x)\,. \ ∎$$

Per questa ragione, $X^{\perp}$ si chiama il COMPLEMENTO ORTOGONALE di X e, reciprocamente, X è il complemento ortogonale di $X^{\perp}$.

Osservazione 225 Grazie a quest'osservazione, la dimostrazione del teorema di Hahn-Banach in spazi di Hilbert si fa in modo elementare. Se L_0 è un funzionale lineare e continuo sul s.spazio X_0 di H, esso si estende prima di tutto per continuità alla chiusura X di X_0 (indichiamo ancora con L_0 tale estensione) e poi il funzionale L_0 si estende ad H definendo

$$L(Px + (x - Px)\,) = L_0 Px\,.$$

Ovviamente, $||L|| = ||L_0||$. Si noti anche che la restrizione di L ad $X^{\perp}$ è nulla. ∎

Proviamo ora:

Teorema 226 *Sia X un s.spazio chiuso dello spazio di Hilbert H. E':*

$$X = \left[X^{\perp}\right]^{\perp}\ .$$

<u>Dim</u>. La (3.8) mostra che $X \subseteq \left[X^{\perp}\right]^{\perp}$. Proviamo l'uguaglianza. Sia $\xi \in \left[X^{\perp}\right]^{\perp}$ e sia ξ_0 la sua proiezione ortogonale su X. Dobbiamo provare che $\xi = \xi_0$.

Sia ξ che ξ_0 sono elementi di $\left[X^{\perp}\right]^{\perp}$ e quindi anche

$$\xi - \xi_0 \in \left[X^{\perp}\right]^{\perp} .$$

Per definizione di proiezione ortogonale, $\xi - \xi_0 \in X^{\perp}$. Dunque si ha

$$\xi - \xi_0 \in \left[X^{\perp}\right]^{\perp} \quad \text{e anche} \quad \xi - \xi_0 \in X^{\perp} .$$

Il Lemma 219, con $A = X^{\perp}$ mostra che $\xi - \xi_0 = 0$ come volevamo. ∎

3.3.1 Proiezioni ortogonali

Alcune proprietà degli operatore di proiezione si sono viste al paragrafo 2.8.1. Ricordiamo: una proiezione in uno spazio di Banach X è un operatore lineare continuo P tale che $P^2 = P$ e si dice che P proietta X su Y se $Y = \mathrm{im}P$. In tal case, $Y = \mathrm{im}P$ è dotato di complementare perché

$$X = (\mathrm{im}P) \oplus (\mathrm{im}(I - P)) \ : \tag{3.10}$$

il complementare è l'immagine dell'operatore $(I - P)$ che a sua volta è una proiezione.

Ovviamente, se Y è dotato di complementare Z allora esiste una proiezione di X su Y. Infatti, ogni $x \in X$ si scrive in modo unico come

$$x = y + z \quad y = y(x) \, , \ z = z(x) \, .$$

La proiezione di X su Y è l'operatore che as x associa $y(x)$. Infatti, si prova che quest'operatore è lineare e continuo. *Però possono esistere s.spazi* chiusi *di uno spazio di Banach, che sono privi di complementari e quindi la proiezione su tali s.spazi non esiste.* Invece, il Teorema 224 mostra che in uno spazio di Hilbert ogni s.spazio chiuso ammette complementare e quindi se X è un qualsiasi s.spazio chiuso di uno *spazio di Hilbert H,* esiste la proiezione di H su X.

Il particolare operatore di proiezione costruito nel Teorema 224 è quello che associa ad ogni elemento $h \in H$ la sua proiezione ortogonale su X. Essa si chiama la PROIEZIONE ORTOGONALE di H su X.

Vale la pena di verificare esplicitamente che le proiezioni ortogonali in uno spazio di Hilbert sono operatori lineari e continui perché nel caso degli spazi di Hilbert la dimostrazione della continuità è elementare e non usa il teorema di Baire.

Teorema 227 *Sia X un s.spazio chiuso di uno spazio di Hilbert H. L'operatore P che ad ogni elemento di H associa la sua proiezione ortogonale su X è una proiezione, ossia un operatore lineare e continuo tale che $P^2 = P$. Inoltre, se $X \neq 0$ si ha $\|P\| = 1$. Se invece $X = 0$ allora $P = 0$.*

Dunque anche $I - P$ è una proiezione e se $X \neq H$ si ha $\|I - P\| = 1$. Se invece $X = H$ allora $I - P = 0$.

Dim. Mostriamo che l'operatore P è lineare e continuo.

Proviamo la linearità. Ricordiamo infatti che $P(\alpha x_0 + \beta x_1)$ è l'unico elemento tale che

$$[(\alpha x_0 + \beta x_1) - P(\alpha x_0 + \beta x_1)] \perp X\,.$$

Ma, $(x_0 - Px_0) \perp X$ e $x_1 - Px_1 \perp X$ così che

$$[(\alpha x_0 + \beta x_1) - (\alpha P x_0 + \beta P x_1)] \perp X\,.$$

Dunque,

$$P(\alpha x_0 + \beta x_1) = \alpha P x_0 + \beta P x_1$$

perché la proiezione ortogonale è unica.

Notare che queste proprietà valgono anche per l'operatore $I - P$ la cui immagine è $X^\perp$.

La continuità di P e anche quella di $I - P$ seguono dal teorema di Pitagora. Infatti, per ogni $h \in H$ vale

$$||h||^2 = ||Ph + (I - P)h||^2 = ||Ph||^2 + ||(I - P)h||^2\,.$$

Dunque,

$$||P|| \leq 1\,, \qquad ||(I - P)|| \leq 1\,.$$

Più precisamente:

$$\begin{cases} \text{if } X \neq 0 \text{ then } & ||P|| = 1\,, \\ \text{if } X \neq H \text{ then } & ||(I - P)|| = 1\,. \end{cases} \tag{3.11}$$

Infatti,

$$\|P\| = \sup_{\|x\|=1} \|Px\| \quad \text{e} \quad Px = x \quad \text{se} \quad x \in X\,.$$

Naturalmente, in questo calcolo si è implicitamente supposto $X \neq 0$, altrimenti $x \in X$ con $\|x\| = 1$ non esiste. Quando $X = 0$ si ha $\|P\| = 0$.

In modo analogo si vede la seconda uguaglianza se $X^\perp \neq \{0\}$, ossia se $X \neq H$.

Le proprietà

$$P^2h = P(Px) = Ph\,, \qquad (I-P)^2h = (I-P)h$$

sono ovvie. ∎

Dal Teorema 227 segue che ogni proiezione ortogonale ha norma 0 oppure 1. Si contrasti con l'Esempio 114.

Due caratterizzazioni importanti delle proiezioni ortogonali sono le seguenti.

Notiamo che

$$\langle Ph, k\rangle = \langle h, Pk\rangle \qquad h\,,\ k \in H\,. \tag{3.12}$$

Infatti,

$$\langle Ph, k\rangle = \langle Ph, Pk + (I-P)k\rangle = \langle Ph, Pk\rangle$$

perché $P(I-P) = 0$. Per questa stessa ragione,

$$\langle h, Pk\rangle = \langle Ph + (I-P)h, Pk\rangle = \langle Ph, Pk\rangle$$

e quindi vale (3.12). ∎

Seguendo la terminologia nota dalla dimensione finita, un operatore lineare continuo per cui vale la (3.12) si dice SIMMETRICO. Dunque, ogni proiezione ortogonale è un operatore simmetrico. Si vede facilmente che vale anche il viceversa:

Teorema 228 *Sia $P \in \mathcal{L}(H)$ una proiezione. L'operatore P è la proiezione* ortogonale *sul s.spazio $X = PH$ se e solo se è simmetrico.*

<u>Dim</u>. La (3.12) mostra che ogni proiezione ortogonale è un operatore simmetrico. Viceversa, mostriamo che se $P \in \mathcal{L}(H)$ è un operatore di proiezione che è anche simmetrico allora P è proiezione ortogonale. Sia per questo $X = \operatorname{im} P$. Mostriamo prima di tutto che X è un s.spazio *chiuso.* Sia per questo (x_n) una successione in X, convergente ad un $h \in H$. Dobbiamo provare che $h \in X$.

Essendo $x_n \in X$, si ha

$$x_n = Px_n\,.$$

Passando al limite, grazie alla continuità di P, si trova

$$h = \lim x_n = \lim Px_n = Ph \in \operatorname{im} P = X\,.$$

Ciò prova che X è chiuso.

Sia ora $h \in H$. Mostriamo che

$$h - Ph \perp X\,,$$

così che Ph è effettivamente la proiezione ortogonale di h su X. Sia per questo x un generico element di X, ossia un generico elemento Pk dell'immagine di P. Si ha

$$\langle h - Ph, x\rangle = \langle h - Ph, Pk\rangle = \langle P(h - Ph), k\rangle = \langle Ph - Ph, k\rangle = 0$$

(si noti che in questo calcolo si è usato il fatto che P è sia una *proiezione* che un operatore *simmetrico*.)

Ciò è quanto volevamo provare. ∎

Proviamo infine:

Teorema 229 *Sia $P \in \mathcal{L}(H)$. L'operazione P è proiezione ortogonale sulla sua immagine X, $X = \mathrm{im}P$, se e solo se*

$$[\mathrm{im}\, P]^{\perp} = \mathrm{im}\,(I - P)\,. \tag{3.13}$$

Dim. Sia P la proiezione ortogonale di H su $X = \mathrm{im}P$. Proviamo che vale l'uguaglianza in (3.13).

Per definizione, $x - Px = (I - P)x \perp \mathrm{im}P$ e quindi

$$\mathrm{im}\,(I - P) \subseteq [\mathrm{im}P]^{\perp} = \{y\ :\ \langle Ph, y\rangle = 0\ \forall h \in H\}\,.$$

L'inclusione opposta vale perché se $\xi \perp \mathrm{im}P$ allora $\xi - 0 \perp \mathrm{im}P$ e quindi ξ ha proiezione nulla su $\mathrm{im}P$. Dunque, $\xi \in \mathrm{im}(I - P)$ da (3.10).

Ciò prova l'uguaglianzain (3.13).

Viceversa, proviamo che se l'uguaglianzain (3.13) vale allora P è proiezione ortogonale di H su $X = \mathrm{im}P$. Ciò vale perché tale uguaglianza significa che

$$h - Ph \perp Pk \qquad \forall k \in H$$

e quindi Ph è la proiezione di H sul s.spazio $\mathrm{im}P$ per la definizione di proiezione ortogonale. ∎

3.3.2 Sistemi ortonormali e calcolo di proiezioni

Un insieme S di vettori di uno spazio di Hilbert si chiama ORTOGONALE se

$$x\,,\ y \in \mathcal{S}\,, \quad x \neq y \Longrightarrow\ x \perp y\,.$$

Se ogni elemento di $\mathcal{S}$ ha norma 1, l'insieme $\mathcal{S}$ si chiama ORTONORMALE.

Ovviamente, *un sistema ortogonale che non contiene* 0 *è linearmente indipendente, e quindi un sistema ortonormale è linearmente indipendente.*

Esponiamo un metodo, detto METODO DI GRAM–SCHMIDT che permette di costruire sistemi ortonormali a partire da un qualsiasi insieme numerabile $X \subseteq H$. Supponiamo per semplicità che $X = \{x_n\}$ sia linearmente indipendente. In tal caso, in particolare, ciascun suo elemento è non nullo.

Associamo a x_1 l'elemento

$$e_1 = \frac{x_1}{||x_1||} .$$

Ad x_2 associamo

$$e_2 = \frac{z_2}{||z_2||} \qquad \text{ove} \qquad z_2 = x_2 - \langle x_2, e_1 \rangle e_1 .$$

Scelti $e_1 ,\dots , e_{n-1}$ definiamo

$$e_n = \frac{z_n}{||z_n||} \qquad \text{ove} \qquad z_n = x_n - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k .$$

E' immediato vedere che gli e_i sono due a due ortogonali ed ovviamente di norma 1. Inoltre,

Teorema 230 *Per ogni n vale*

$$\text{span}\,\{e_1 , \dots , e_n\} = \text{span}\,\{x_1 , \dots , x_n\} .$$

Osservazione 231 Abbiamo visto che la sfera di uno spazio di Banach di dimensione infinita non è compatta. Ovviamente ciò vale in particolare per gli spazi di Hilbert. Però nel caso degli spazi di Hilbert si può dare una dimostrazione elementare: col metodo precedente si costruisce un sistema numerabile ed ortonormale $\{e_n\}$. Si nota quindi che la successione (e_n) non ha s.successioni convergenti. Infatti, per $n \neq m$ si ha

$$||e_n - e_m||^2 = 2 . \quad \blacksquare$$

Mostriamo ora come i sistemi ortonormali numerabili si possano usare per il calcolo di proiezioni. Consideriamo prima di tutto il caso in cui X sia un s.spazio di H, *di dimensione finita k.*

Sia

$$e_1 , \ \dots , \ e_k$$

una BASE ORTONORMALE di X.

In tal caso la proiezione x_0 di h su X è data da

$$x_0 = \sum_{i=1}^{k} \alpha_i e_i$$

perché ogni elemento di X ha questa forma.

I numeri α_i si calcolano facilmente con un calcolo del tutto analogo a quello noto quando lo spazio H ha dimensione finita:

$$\langle h, e_r \rangle = \langle \sum_{i=1}^{k} \alpha_i e_i, e_r \rangle = \alpha_r \,,$$

Dunque,

$$x_0 = \sum_{i=1}^{k} \langle h, e_i \rangle e_i \,.$$

E' utile calcolare ora

$$||x_0||^2 = \langle \sum_{i=1}^{k} \alpha_i e_i, \sum_{j=1}^{k} \alpha_j e_j \rangle = \sum_{i,j=1}^{n} \alpha_i \bar{\alpha}_j \langle e_i, e_j \rangle = \sum_{i=1}^{n} |\alpha_i|^2$$

perché i vettori e_i sono due a due ortogonali e di norma 1.

Ricordando l'espressione di α_i e la (3.11) si trova

$$\sum_{i=1}^{n} |\langle h, e_i \rangle|^2 = ||x_0||^2 \le ||h||^2 \,. \tag{3.14}$$

Sia ora $S = \{e_i\}$ un sistema ortonormale numerabile. Il s.spazio di H

$$\operatorname{span} S = \left\{ \sum_{i=1}^{n} \alpha_i e_i \ \ \alpha_i \in \mathbb{C} \,, \ n \in \mathbb{N} \right\}$$

non è chiuso. Indichiamo con X la sua chiusura. Vogliamo rappresentare x_0, la proiezione su X di un generico elemento $h \in H$.

Notiamo prima di tutto:

Lemma 232 *Supponiamo che (e_i) sia una successione ortonormale e che la serie $\sum_{i=1}^{+\infty} \alpha_i e_i$ converga in H. Vale*

$$\left|\left| \sum_{i=1}^{+\infty} \alpha_i e_i \right|\right|^2 = \sum_{i=1}^{+\infty} |\alpha_i|^2 \,. \tag{3.15}$$

In particolare se la serie $\sum_{i=1}^{+\infty} \alpha_i e_i$ converge in H allore $(\alpha_i) \in l^2$.

Dim. La somma di una serie è il limite della successione delle somme parziali,

$$\sum_{i=1}^{+\infty} \alpha_i e_i = \lim_n \sum_{i=1}^{n} \alpha_i e_i$$

e, per la continuità della norma,

$$\left\| \sum_{i=1}^{+\infty} \alpha_i e_i \right\|^2 = \left\| \lim_n \sum_{i=1}^{n} \alpha_i e_i \right\|^2 = \lim_n \left\| \sum_{i=1}^{n} \alpha_i e_i \right\|^2 .$$

L'asserto segue dall'uguaglianza

$$\left\| \sum_{i=1}^{n} \alpha_i e_i \right\|^2 = \sum_{i=1}^{n} |\alpha_i|^2 . \ \blacksquare$$

Completando questo risultato con l'implicazione opposta si trova il TEOREMA DI RIESZ–FISCHER:

Teorema 233 (di RIESZ–FISCHER) *Sia $\mathcal{S} = \{e_i\}$ un sistema ortonormale numerabile. La serie*

$$\sum_{i=1}^{n} \alpha_i e_i$$

converge in H se e solo se la successione (α_i) è in l^2.

Dim. Dal Lemma 232, se la serie converge in H si ha

$$\sum_{i=1}^{+\infty} |\alpha_i|^2 = \left\| \sum_{i=1}^{+\infty} \alpha_i e_i \right\|^2 < +\infty .$$

Il viceversa segue notando che se $(\alpha_n) \in l^2$ allora la successione delle somme parziali è fondamentale. Infatti,

$$\left\| \sum_{i=n}^{m} \alpha_i e_i \right\|^2 = \sum_{i=n}^{m} |\alpha_i|^2$$

e la serie $\sum_{i=1}^{+\infty} |\alpha_i|^2$ converge perché per ipotesi la successione (α_i) è in l^2. $\blacksquare$

Indichiamo ora con X_n lo spazio lineare (di dimensione finita) generato dai vettori $e_1, \ldots, e_n$. Come si è visto, la proiezione x_n di h su X_n è

$$x_n = \sum_{i=1}^{n} \langle h, e_i \rangle e_i$$

e, dalla (3.14), per ogni n vale

$$\sum_{i=1}^{n} |\alpha_i|^2 = \sum_{i=1}^{n} |\langle h, e_i\rangle|^2 \leq ||h||^2 .$$

Dunque la successione (α_i) è in l^2 e quindi

$$x_0 = \sum_{i=1}^{+\infty} \langle h, e_i\rangle e_i \tag{3.16}$$

converge in H. E' facile immaginare che valga:

Teorema 234 *Il vettore x_0 definito in (3.16) è la proiezione ortogonale di h su X.*

Dim. Per mostrare ciò si prova che $h - x_0$ è ortogonale ad ogni elemento di X. Ricordiamo che per definizione ogni $x \in X$ è limite di una successione (s_n) in $\operatorname{span}\mathcal{S}$ e, per la continuità del prodotto interno,

$$\langle x, h - x_0\rangle = \langle \lim s_n, h - x_0\rangle = \lim\langle s_n, h - x_0\rangle .$$

Dunque basta provare che $h - x_0$ è ortogonale a $\operatorname{span}\mathcal{S}$ e per questo basta provare che è ortogonale ad ogni elemento e_k. Ciò si vede immediatamente perchè[3]

$$\langle h - x_0, e_k\rangle = \langle h, e_k\rangle - \langle \sum_{i=1}^{+\infty} \langle h, e_i\rangle e_i, e_k\rangle$$

$$= \langle h, e_k\rangle - \sum_{i=1}^{+\infty} \langle\langle h, e_i\rangle e_i, e_k\rangle = \langle h, e_k\rangle - \langle h, e_k\rangle = 0 . \quad \blacksquare$$

Abbiamo così identificato la proiezione x_0 di h su X,

$$x_0 = \sum_{i=1}^{+\infty} \langle h, e_i\rangle e_i .$$

Dalla (3.14), vale

$$\sum_{i=1}^{+\infty} |\langle h, e_i\rangle|^2 \leq ||h||^2 .$$

Questa disuguaglianza si chiama DISUGUAGLIANZA DI BESSEL.

In particolare,

[3] si noti l'uso della linearità e continuità della prima componente del prodotto interno, per scambiare i segni di serie e di prodotto interno.

Teorema 235 *Ogni x_0 che appartiene all'insieme*

$$\operatorname{cl}\operatorname{span} S = \operatorname{cl}\left\{\sum_{i=1}^{n} \alpha_i e_i \ \ \alpha_i \in \mathbb{C}\,,\ n \in \mathbb{N}\right\}$$

ha una unica rappresentazione in serie degli elementi e_i

$$x_0 = \sum_{i=1}^{+\infty} \langle x_0, e_i\rangle e_i \,.$$

Per questi elementi x_0 la disuguaglianza di Bessel vale col segno di uguale:

$$\sum_{i=1}^{+\infty} |\langle x_0, e_i\rangle|^2 = \left\|\sum_{i=1}^{+\infty}\langle x_0, e_i\rangle e_i\right\|^2 \,.$$

Grazie a questo risultato, la successione $\{e_i\}$ si chiama BASE ORTONORMALE dello spazio $\operatorname{cl}\operatorname{span} S$ da essa generato

3.3.3 Serie di Fourier astratte

Se H è uno spazio di Hilbert separabile e di dimensione infinita. Usando il Lemma di Zorn si può provare che in H esistono (infinite) successioni (e_n) tali che

- L'insieme $S = \{e_n\}$ è ortonormale;
- lo spazio lineare (non chiuso) $\operatorname{span}\{e_n\}$ è denso in H.

In tal caso $\mathcal{S}$ si chiama un SISTEMA ORTONORMALE MASSIMALE O COMPLETO)[4]. L'insieme $\mathcal{S}$ si chiama anche BASE ORTONORMALE di H.

E' utile conoscere alcuni test utili per verificare se un sistema ortonormale numerabile in uno spazio di Hilbert è massimale o meno. Vale:

Teorema 236 *Sia $\mathcal{S} = \{e_i\}$ un sistema ortonormale finito o numerabile in uno spazio di Hilbert H. Si equivalgono le affermazioni seguenti:*

i) *il sistema $\mathcal{S}$ è massimale;*

ii) *ogni $h \in H$ si sviluppa in serie degli e_i,*

$$h = \sum \alpha_i e_i \qquad \alpha_i = \langle h, e_i\rangle\,; \tag{3.17}$$

[4]in realtà vale di più: usando il lemma di Zorn, si prova che *ogni spazio di Hilbert ha un sistema ortonormale massimale, e che questo è numerabile se e solo se H è separabile.*

iii) *per ogni $h \in H$ vale l'uguaglianza*

$$||h||^2 = \sum_{i=1}^{+\infty} |\langle h, e_i \rangle|^2 \ ; \tag{3.18}$$

iv) *se $\langle h, e_i \rangle = 0$ per ogni i allora $h = 0$.*

Dim. Si è già visto che **i)** implica **ii)** e quindi **iii)** vale per il lemma 232. In particolare, se $\langle h, e_i \rangle = 0$ per ogni i allora $h = 0$, ossia vale **iv)**. La dimostrazione si completa provando che se vale **iv)** allora $\mathcal{S}$ è massimale.

La condizione **iv)** significa

$$[\mathrm{span}\,\mathcal{S}]^\perp = \{0\}$$

Si sa, dal teorema 221 che in tal caso $\mathrm{span}\,\mathcal{S}$ è denso in H. Dunque, $\mathcal{S}$ è massimale. ∎

L'uguaglianza (3.18) si chiama IDENTITÀ DI PARSEVAL e la rappresentazione (3.17) si chiama SERIE DI FOURIER ASTRATTA.

3.4 Esempi di sistemi ortogonali massimali

In questa parte consideriamo esplicitamente alcuni esempi di sviluppi in serie di funzioni ortogonali che sono maggiormente importanti per le applicazioni. Ovviamente, il più importante è quello delle SERIE DI FOURIER, ma altri sono gli sviluppi in serie di POLINOMI ORTOGONALI e gli sviluppi in serie di FOURIER-BESSEL.

3.4.1 Sviluppi in serie di Fourier

Consideriamo lo spazio di Hilbert $L^2(-\pi, \pi)$. Una verifica immediata è che il sistema di funzioni

$$\frac{1}{\sqrt{2\pi}} e^{inx} \tag{3.19}$$

è un sistema ortonormale. Vogliamo provare

Teorema 237 *Il sistema (3.19), al variare di n tra i numeri* interi, *è completo in $L^2(-\pi, \pi)$.*

Dim. E' sufficiente mostrare che lo spazio lineare generato da queste funzioni è denso in $L^2(-\pi,\pi)$ e per questo basta mostrare che ogni funzione di un opportuno s.spazio denso in $L^2(-\pi,\pi)$ è approssimabile in norma quadratica mediante combinazioni lineari delle funzioni in (3.19).

Un s.spazio denso in $L^2(-\pi,\pi)$ si è già trovato, ed è quello dei polinomi, si veda il Teorema 49.

E' immediatamente evidente che ogni polinomio può approssimarsi nella norma di $L^2(-\pi,\pi)$ mediante una successione di funzioni $f(x)$ con queste proprietà:

- la funzione $f(x)$ è di classe C^2;
- vale $f(-\pi)=f(\pi)$, $f'(-\pi)=f'(\pi)$, $f''(-\pi)=f''(\pi)$.

Chiameremo queste le funzioni "periodiche di classe C^2" perché si possono estendere a funzioni periodiche (definite su $\mathbb{R}$) e di classe C^2. Parleremo di "funzioni periodiche di classe C^1" se la condizione sulla derivata seconda non viene richiesta.

Ricapitolando, l'insieme delle funzioni periodiche di classe C^2 è denso in $L^2(-\pi,\pi)$.

Proveremo che ogni funzione periodica di classe C^2 è limite uniforme (e quindi anche in norma $L^2(-\pi,\pi)$) della propria serie di Fourier:

$$f(x)=\lim_N s_N(x)\,,\qquad s_N(x)=\sum_{n=-N}^{+N} f_n e^{inx} \tag{3.20}$$

dove

$$f_n=\frac{1}{2\pi}\int_{-\pi}^{+\pi} f(\xi)e^{-in\xi}\,\mathrm{d}\xi\,.$$

Ciò è sufficiente per provare il teorema.

Servono per questo tre osservazioni:

- per ogni $N\geq 0$ vale

  $$\frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{n=-N}^{N} e^{in(x-\xi)}\,\mathrm{d}\xi = 1+\sum_{n=-N;\ n\neq 0}^{N}\frac{1}{in}\left[e^{in(x+\pi)}-e^{in(x-\pi)}\right]=1$$

  perché

  $$e^{in(x-\pi)}=e^{in(x+\pi)}\,.$$

- se $\phi_x(\xi)$ è una famiglia di funzioni periodiche di classe C^1, dipendente dal parametro x e con derivate (rispetto a ξ) limitate uniformemente al variare di x,

  $$|\phi_x'(\xi)| < M \qquad \forall x\,,\ \xi$$

  allora[5]

  $$\lim_n \int_{-\pi}^{\pi} \phi_x(\xi) e^{in(x-\xi)}\,\mathrm{d}\xi = 0\,. \tag{3.21}$$

  Il limite è uniforme rispetto ad x.

  Questa proprietà si vede notando che

  $$e^{in(x-\xi)} = (i/n)\frac{\mathrm{d}}{\mathrm{d}\xi}(e^{in(x-\xi)})\,.$$

  Integrando per parti:

  $$\left|\int_{-\pi}^{\pi} \phi_x(\xi)e^{in(x-\xi)}\,\mathrm{d}\xi\right| = \left|\frac{i}{n}\left[\phi_x(\xi)e^{in(x-\xi)}\right]_{\xi=-\pi}^{+\pi} + \frac{i}{n}\int_{-\pi}^{\pi} \phi_x'(\xi)e^{in(x-\xi)}\,\mathrm{d}\xi\right| = \frac{1}{n}\left|\int_{-\pi}^{\pi} \phi_x'(\xi)e^{in(x-\xi)}\,\mathrm{d}\xi\right| < \frac{2\pi M}{n} \to 0$$

  per $n \to +\infty$. In particolare, prendendo le parti reali ed immaginarie in (3.21), si vede che vale anche

  $$\begin{cases} \lim_n \int_{-\pi}^{\pi} \phi_x(\xi)\sin[n(x-\xi)]\,\mathrm{d}\xi = 0\,, \\ \lim_n \int_{-\pi}^{\pi} \phi_x(\xi)\cos[n(x-\xi)]\,\mathrm{d}\xi = 0 \end{cases} \tag{3.22}$$

  e il limite è uniforme rispetto ad x.

- Vale[6]

  $$D_N(x) = \sum_{k=-N}^{N} e^{ikx} = \begin{cases} 2N+1 & \begin{cases} \text{se } x = 0 \\ \text{oppure } x = 2\pi \end{cases} \\ \frac{\sin\left[\frac{2N+1}{2}x\right]}{\sin(x/2)} & \text{altrimenti.} \end{cases} \tag{3.23}$$

---

[5]Si noti che questa è una versione semplicissima del Teorema di Riemann-Lebesgue, Teorema 379.

[6]si veda la fine della dimostrazione del Teorema 116.

Proviamo ora la (3.20). Ricordiamo che $f(x)$ è periodica di classe C^2 e scriviamo

$$
\begin{aligned}
& f(x) - s_N(x) = f(x)\cdot 1 - s_N(x) \\
& = \frac{1}{2\pi}\int_{-\pi}^{+\pi} [f(x)-f(\xi)] \sum_{n=-N}^{+N} \left[e^{-in\xi}e^{inx}\right] \,\mathrm{d}\xi \\
& = \frac{1}{2\pi}\int_{-\pi}^{+\pi} [f(x)-f(\xi)] \frac{\sin[(2N+1)(x-\xi)/2]}{\sin[(x-\xi)/2]} \,\mathrm{d}\xi = \\
& \frac{1}{2\pi}\int_{-\pi}^{+\pi} \sin[(N+1/2)(x-\xi)] \frac{f(x)-f(\xi)}{\sin[(x-\xi)/2]} \,\mathrm{d}\xi \\
& = \frac{1}{2\pi}\int_{-\pi}^{\pi} [\sin N(x-\xi)] \frac{f(x)-f(\xi)}{\tan[(x-\xi)/2]} \,\mathrm{d}\xi \qquad (3.24) \\
& + \frac{1}{2\pi}\int_{-\pi}^{\pi} [\cos N(x-\xi)] \left\{f(x)-f(\xi)\right\} \mathrm{d}\xi \,. \qquad (3.25)
\end{aligned}
$$

L'integrale in (3.25) tende a zero per $N \to +\infty$, grazie a (3.22) con $\phi_x(\xi) = f(x) - f(\xi)$. Mostriamo che la proprietà in (3.22) può applicarsi anche all'integrale in (3.24). Essendo $\tan t$ periodica di periodo π, ciascuna delle funzioni

$$\xi \to \phi_x(\xi) = \frac{f(x)-f(\xi)}{\tan[(x-\xi)/2]}$$

è periodica di periodo 2π. Proviamo che la funzione $\xi \to \phi_x(\xi)$ è periodica di classe C^1 con derivata uniformemente limitata rispetto ad x e ξ, così che l'addendo in (3.24) tende a zero per $N \to +\infty$.

Essendo $f(\xi)$ differenziabile, $f(x) - f(\xi) = \mathrm{o}(x-\xi)$, si ha che $\xi \to \phi_x(\xi)$ ammette estensione continua a $\xi = x$. Mostriamo che la sua derivata è periodica e continua anche per $\xi = x$, e che c'è una limitazione della derivata uniforme rispetto ad x e ξ, grazie al fatto che $f(\xi)$ ammette *due* derivate continue.

La derivata $\frac{\mathrm{d}}{\mathrm{d}\xi}\phi_x(\xi)$ è:

$$
\begin{aligned}
& \frac{-f'(\xi)\tan[(x-\xi)/2] + (1/2)[f(x)-f(\xi)]\left\{1/\cos^2[(x-\xi)/2]\right\}}{\tan^2[(x-\xi)/2]} \\
& = \frac{f(x)-f(\xi)-f'(\xi)\sin(x-\xi)}{2\sin^2[(x-\xi)/2]} \,.
\end{aligned}
$$

Questa funzione è periodica di periodo 2π. Per mostrare la regolarità si scrive il numeratore come

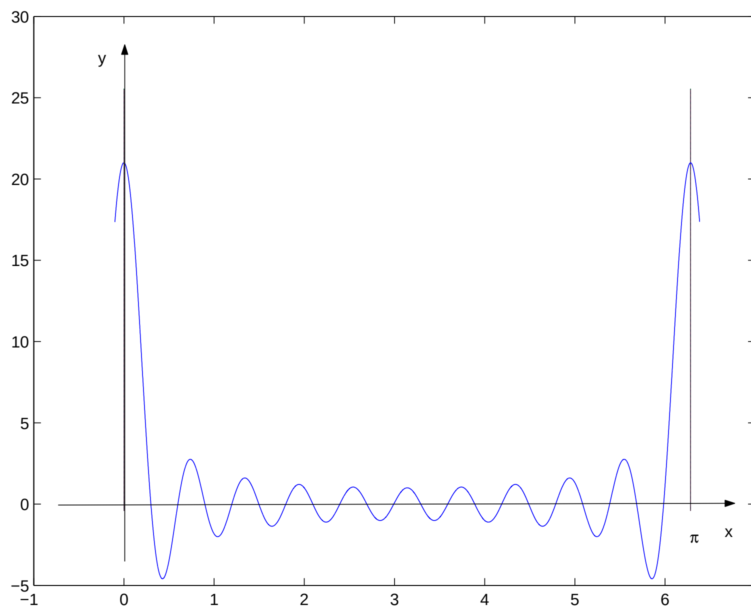
$$\left\{f(x)-f(\xi)-f'(\xi)(x-\xi)\right\} + f'(\xi)\left\{(x-\xi)-\sin(x-\xi)\right\} \,.$$

La limitatezza della derivata $(\mathrm{d}/\mathrm{d}\xi)\phi_x(\xi)$ segue usando la continuità della derivata *seconda* di $f(\xi)$ e scrivendo la formula di Taylor arrestata al secondo ordine con resto in forma di Lagrange di ciascuno dei termini in parentesi graffa.

Osservazione 238 Abbiamo provato che le somme parziali delle serie di Fourier di $f(x)$ convergono uniformemente ad $f(x)$ se $f(x)$ è periodica e di classe C^2. E' facile immaginare che le condizioni di regolarità possano indebolirsi, e si potrebbe congetturare che sia sufficiente la sola continuità di $f(x)$. Si sa, dal teorema 116, che questa congettura è falsa.

Come si è detto, la successione di funzioni $(D_k(x))$ si chiama NUCLEO DI DIRICHLET di Dirichlet e non è un'identità approssimata (il grafico di $D_{10}(x)$ è in figura (3.1)) dato che le funzioni $D_N(x)$ non hanno segno costante. Ciò mostra che la convergenza della serie di Fourier dipende dalla compensazione di valori negativi e positivi e quindi è un fenomeno piuttosto delicato. ∎

Figura 3.1:



Gli sviluppi in serie di Fourier sono i primi esempi di sviluppi in serie rispetto a sistemi ortogonali completi che sono stati studiati, e sono ancora

l'esempio più importante per le applicazioni. Per questa ragione, quando $\{e_i\}$ è un sistema ortonormale massimale in uno spazio di Hilbert H, la serie

$$h = \sum_{i=1}^{+\infty} \langle h, e_i \rangle e_i$$

si chiama SERIE DI FOURIER ASTRATTA.

3.4.2 Una conseguenza dell'identità di Parseval

Sia $f(x)$ una funzione a quadrato integrabile su $[-\pi, \pi]$. La funzione della variabile reale ω definita da

$$\hat{f}(\omega) = \int_{-\pi}^{\pi} e^{-i\omega t} f(t)\, \mathrm{d}t$$

si chiama[7] TRASFORMATA DI FOURIER di f.

Consideriamo ora la base ortonormale $\{e_n\}$ con

$$e_n = \frac{1}{\sqrt{2\pi}} e^{int} .$$

Allora,

$$\hat{f}(n) = \sqrt{2\pi} \langle f, e_n \rangle .$$

Dunque, dall'identità di Parseval, si trova:

$$||f||^2_{L^2(-\pi,\pi)} = \frac{1}{2\pi} \sum_{n=-\infty}^{n=+\infty} |\hat{f}(n)|^2 . \tag{3.26}$$

Quest'espressione si può scrivere in un modo più generale. Fissiamo un qualsiasi numero reale T e consideriamo la funzione

$$g(t) = e^{-iTt} f(t) .$$

Si calcola facilmente che

$$\hat{g}(\omega) = \hat{f}(\omega + T) .$$

Leggiamo questa con $\omega = n$, numero intero, e notiamo che se T è reale si ha

$$||f||_{L^2(-\pi,\pi)} = ||g||_{L^2(-\pi,\pi)} .$$

[7] vedremo più avanti che questa definizione è un caso particolare di una definizione più generale.

D'altra parte, anche per g vale

$$\langle g, e_n\rangle = \frac{1}{\sqrt{2\pi}}\hat{g}(n)\,.$$

Applicando l'identità (3.26) alla funzione g si trova:

$$||f||^2_{L^2(-\pi,\pi)} = ||g||^2_{L^2(-\pi,\pi)} = \frac{1}{2\pi}\sum_{n=-\infty}^{n=+\infty} |\hat{f}(n+T)|^2\,.$$

Si noti che quest'uguaglianza vale per ogni T e quindi la somma della serie non *dipende da T.*

3.4.3 Polinomi di Legendre, Laguerre ed Hermite

Senza ancora specificare lo spazio in cui si lavora, consideriamo l'operatore che ad una funzione $y(x)$ associa

$$(\mathcal{A}y)\,(x) = \sigma(x)y''(x) + \tau(x)y'(x)\,. \tag{3.27}$$

Accade frequentemente di dover calcolare numeri λ in corrispondenza dei quali esistono *polinomi* $y(x)$ che risolvono l'equazione

$$\mathcal{A}y = \lambda y$$

ossia

$$\sigma(x)y''(x) + \tau(x)y'(x) = \lambda y(x) \tag{3.28}$$

Ovviamente, dopo aver specificato lo spazio in cui si intende lavorare, questo sarà un problema di calcolo di autovalori ed autofunzioni.

Osservazione 239 Quando si lavora con equazioni differenziali, è pratica comune non indicare esplicitamente la variabile indipendente della funzione incognita. Noi seguiremo questa pratica e quindi, per esempio, invece di (3.28), scriveremo

$$\sigma(x)y'' + \tau(x)y' = \lambda y\,. \quad \blacksquare$$

Notiamo subito una forma diversa sotto cui si può porre questo problema: sia $\rho(x)$ soluzione non nulla di

$$(\sigma(x)\rho)' = \tau(x)\rho\,.$$

Tabella 3.1: Polinomi ortogonali classici. Si noti che il peso $\rho(x)$ è definito a meno di un fattore costante e positivo.

	Legendre: $P_n(x)$	Laguerre: $L_n(x)$	Hermite: $H_n(x)$
$\sigma(x)$	$(1-x^2)$	x	1
$\tau(x)$	$-2x$	$-x+1$	$-2x$
$\rho(x)$	1	e^{-x}	e^{-x^2}
λ_n	$-n(n+1)$	$-n$	$-2n$
H	$L^2(-1,1)$	$L^2(0,+\infty;\rho)$	$L^2(-\infty,+\infty;\rho)$

Moltiplicando per ρ i due membri di (3.28) si vede che risolvere la (3.28) equivale a risolvere

$$(\rho(x)\sigma(x)y')' = \lambda\rho(x)y\,. \tag{3.29}$$

Si noti inoltre che $\rho(x)$, essendo una soluzione *non nulla* di un'equazione differenziale lineare del primo ordine, non ha zeri. Non è quindi restrittivo assumere

$$\rho(x) \geq 0\,,$$

come da ora in poi faremo.

Diciamo subito che troveremo successioni di autovalori e successioni di autofunzioni che sono *polinomi*, in uno spazio che ora specificheremo. Con le scelte particolari di $\sigma(x)$ e di $\tau(x)$ che ora vedremo, questi si chiamano i POLINOMI ORTOGONALI CLASSICI. E precisamente, con le scelte particolari nella Tabella 3.1, si parla di POLINOMI DI LEGENDRE, POLINOMI DI LAGUERRE e POLINOMI DI HERMITE. La tabella riporta $\sigma(x)$, $\tau(x)$, $\rho(x)$ e gli autovalori nei tre casi. La prima riga riporta i simboli standard che si usano per indicare tali polinomi e l'ultima riga riporta lo spazio di Hilbert in cui, come vedremo, questi sono un sistema ortogonale completo.

Le informazioni nella tabella 3.1 sono ripetute, e completate, nella tabella 3.2.

Osservazione 240 Si osservi che gli autovalori sono *negativi.* Talvolta invece di usare l'espressione (3.28), si studia il problema

$$\sigma(x)y'' + \tau(x)y' + \lambda y = 0\,. \quad \blacksquare$$

In questo paragrafo studieremo le proprietà principali dei polinomi ortogonali classici, in modo unificato. Per questo indicheremo genericamente con $p_n(x)$ uno di tali polinomi, che leggeremo nello spazio di Hilbert detto nella tabella, e che indicheremo con H.

Osservazione 241 La proprietà principale delle funzioni $\sigma(x)$ e $\tau(x)$ è di essere polinomi rispettivamente di grado al più 2 la $\sigma(x)$ e di grado 1 la $\tau(x)$ e tali che *se $p(x)$ ha grado n allora $(\mathcal{A}p)(x)$ ha ancora grado n.* Inoltre, in ciascuno dei tre casi, $\tau'(0) \equiv \tau'(0) < 0$.

La proprietà principale degli spazi H è che essi contengono tutti i polinomi. Anzi, una dimostrazione analoga a quella del Teorema 49 mostra che *l'insieme dei polinomi è denso in H.* ∎

Come si è detto, vogliamo provare l'esistenza di autovalori e di autovettori che sono *polinomi.* Introduciamo per questo lo spazio lineare $\mathcal{P}_n$ dei polinomi di grado n al più. La dimensione di $\mathcal{P}_n$ è $n+1$ e, come si è notato nell'osservazione 241, l'operatore $\mathcal{A}$ trasforma ciascuno spazio $\mathcal{P}_n$ in sé. Essendo lo spazio $\mathcal{P}_n$ di dimensione finita, l'operatore $\mathcal{A}$ ammette almeno un autovalore λ ed un autovettore, che è quindi un polinomio di grado n al più, in ciascuno degli spazi $\mathcal{P}_n$.

Notiamo che un autovettore che appartiene a $\mathcal{P}_n$ è anche un autovettore che appartiene a $\mathcal{P}_k$ per ogni $k > n$.

Consideriamo ora lo spazio $\mathcal{P}_0$, i cui elementi sono le costanti. E' chiaro che ciscuna di esse è un autovettore di $\mathcal{A}$, di autovalore uguale a 0.

Consideriamo ora lo spazio $\mathcal{P}_1$ i cui elementi sono i polinomi delle forma $ax + b$ con a e b reali. Mostriamo che in $\mathcal{P}_1$ esistono *due* autovettori di $\mathcal{A}$, di gradi 0 e 1.

L'esistenza di autovettori di grado 0 si è già notata: sono i polinomi costanti.

Notiamo che se si vuole che $y(x)$ risolva (3.28) e se inoltre $y(x)$ deve essere un polinomio di grado 1, allora $y(x)$ risolve

$$\tau(x)y' = \lambda y\,.$$

Mostriamo che quest'equazione, per un opportuno valore di λ, ha effettivamente una soluzione di forma

$$y(x) = x + b\,.$$

Ricordiamo per questo che $\tau'(0) \neq 0$.

Sostituendo ed uguagliando i coefficienti dello stesso grado si trova che λ e b sono dati da

$$\begin{cases} \lambda = \tau'(0) \\ b = \frac{\tau(0)}{\lambda} = \frac{\tau(0)}{\tau'(0)} \,. \end{cases}$$

Si noti che $\lambda < 0$ perché $\tau'(0) < 0$.

Ricapitolando, abbiamo trovato due polinomi $p_0(x)$, di grado 0, e $p_1(x)$, di grado 1, che appartengono a $\mathcal{P}_1$ e che sono autovettori di $\mathcal{A}$. Altri autovettori da essi linearmente indipendenti non esistono perché $\mathcal{P}_1$ ha dimensione 2.

Quest'osservazione si può generalizzare procedendo per induzione.

Supponiamo di aver trovato autovalori

$$\lambda_0 = 0,\ \lambda_1 = \tau'(0),\ \lambda_2,\ \dots,\ \lambda_n$$

e corrispondenti autovettori

$$p_0(x),\ p_1(x),\ p_2(x),\ \dots,\ p_n(x)$$

in $\mathcal{P}_n$. Supponiamo inoltre che il polinomio $p_k(x)$ abbia grado k e che l'espressione per λ_k sia

$$\lambda_k = k\left[\frac{\sigma''(0)}{2}(k-1) + \tau'(0)\right] .$$

Proviamo che si trova un numero λ_{n+1} e un polinomio $p_{n+1}(x)$ di grado $n+1$ e tali che

$$\left(\mathcal{A}p_{n+1}\right)(x) = \lambda_{n+1}p_{n+1}(x)\,, \qquad \lambda_{n+1} = (n+1)\left[\frac{\sigma''(0)}{2}n + \tau'(0)\right] .$$

Si noti che ogni polinomio $r(x)$ di grado n al più, può rappresentarsi come combinazione lineare dei $p_k(x)$, con $k \le n$, perché $p_k(x)$ ha grado k.

Costruiamo un ulteriore autovettore $q(x)$ di grado $n+1$. Ricerchiamlo di forma

$$q(x) = x^{n+1} + \sum_{j=0}^{n} \alpha_j p_j(x)\,.$$

Notiamo che il polinomio x^{n+1} viene trasformato da $\mathcal{A}$ nel polinomio

$$\begin{aligned} &(n+1)n\sigma(x)x^{n-1} + (n+1)\tau(x)x^n \\ &= (n+1)n\left[\sigma(0) + x\sigma'(0) + \frac{x^2}{2}\sigma''(0)\right]x^{n-1} + (n+1)\left[\tau(0) + x\tau'(0)\right]x^n \,. \end{aligned}$$

Imponendo
$$\mathcal{A}q(x) = \mu q(x)$$
si trova
$$(n+1)n\left[\sigma(0) + x\sigma'(0) + \frac{x^2}{2}\sigma''(0)\right]x^{n-1}$$
$$+(n+1)\left[\tau(0) + x\tau'(0)\right]x^n + \sum_{j=0}^{n}\lambda_j\alpha_j p_j(x) = \mu x^{n+1} + \sum_{j=0}^{n}\mu\alpha_j p_j(x)\,.$$

Uguagliando i coefficienti dei termini di grado massimo nei due membri, si trova
$$\mu = (n+1)\left[\frac{\sigma''(0)}{2}n + \tau'(0)\right]\,.$$

Si noti che μ non coincide con nessuno dei λ_k, $k \leq n$.

A questo punto, per verificare l'uguaglianza, basta determinare i coefficienti α_j in modo da avere
$$\sum_{j=0}^{n}[\mu - \lambda_j]\alpha_j p_j(x) = (n+1)n[\sigma(0) + \sigma'(0)x]x^{n-1} + (n+1)\tau(0)x^n\,.$$

Ciò può farsi perché il grado di $p_j(x)$ è j e inoltre $[\mu - \lambda_j] \neq 0$ per ogni j così che $\{(\mu - \lambda_j)p_j(x)\}$ è un sistema linearmente indipendente in $\mathcal{P}_n$.

Osservazione 242 Si noti quindi che $\mathcal{A}$ ha esattamente $n+1$ autovettori distinti nello spazio $\mathcal{P}_n$, di dimensione $n+1$.

Nel calcolo precedente abbiamo anche specificato l'espressione dei λ_n. Ci si potrebbe chiedere se esistano valori λ diversi da questi in corrispondenza dei quali la (3.27) ammetta soluzioni polinomiali. La risposta è negativa: se ciò accadesse, per un certo valore di n, l'operatore $\mathcal{A}$ avrebbe in $\mathcal{P}_n$ più di $n+1$ autovettori linearmente indipendenti e ciò non può aversi perché lo spazio ha dimensione $n+1$.

Il polinomio $p_n(x)$ è soluzione dell'equazione (3.28) con $\lambda = \lambda_n$. Essendo questa un'equazione differenziale lineare del secondo ordine, essa ammette un seconda soluzione, linearmente indipendente da $p_n(x)$. Per quanto detto sopra, questa non è un polinomio e si potrebbe provare che questa funzione *non* appartiene ad H. ∎

Mostriamo ora:

Teorema 243 *La successione $\{p_n(x)\}$ è un sistema ortogonale massimale in H.*

Dim. Indichiamo con (a,b) l'intervallo, eventualmente illimitato, su cui le funzioni di H si considerano, si veda la Tabella 3.1. Proviamo che se $n \neq m$ allora

$$\langle p_n, p_m \rangle_H = \int_a^b p_n(x)\rho(x)\overline{p_m}(x)\,\mathrm{d}x = 0\,.$$

Ricordiamo che $\lambda_n \neq \lambda_m$ e quindi, da

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\sigma(x)\rho(x)p_n'\right] = \lambda_n\rho(x)p_n\,, \qquad \frac{\mathrm{d}}{\mathrm{d}x}\left[\sigma(x)\rho(x)p_m'\right] = \lambda_m\rho(x)p_m\,,$$

si trova

$$\begin{aligned}
&(\lambda_n - \lambda_m)\langle p_n, p_m \rangle_H \\
&= \int_a^b \left\{ \frac{\mathrm{d}}{\mathrm{d}x}\left[\sigma(x)\rho(x)p_n'(x)\right]\overline{p_m}(x) - p_n(x)\overline{\frac{\mathrm{d}}{\mathrm{d}x}\left[\sigma(x)\rho(x)p_m'(x)\right]} \right\} \mathrm{d}x \\
&[\sigma(x)\rho(x)p_n'(x)]\overline{p_m}(x)\Big|_a^b - p_n(x)[\sigma(x)\rho(x)\overline{p_m'}(x)]\Big|_a^b = 0
\end{aligned}$$

tenendo conto dei limiti di $\rho(x)$ e di $\sigma(x)$ per x tendente agli estremi di (a,b).

Come si è detto, ogni polinomio è combinazione lineare dei $p_n(x)$; e l'insieme di tutti i polinomi è denso in H, si veda l'osservazione 241. Ciò prova che il sistema ortonormale dei $p_n(x)$ è massimale in H. ∎

3.4.3.1 Sviluppi in serie di polinomi ortogonali

Dividendo ciascuno dei polinomi $p_n(x)$ per la sua norma nel corrispondente spazio H, si trova una successione ortonormale massimale in H; e quindi ogni $f \in H$ si rappresenta come

$$f(x) = \sum_{n=0}^{+\infty} \left[\langle f, p_n \rangle_H\right] p_n(x)$$

dove il prodotto interno è quello di H e la serie converge nella norma di H. Tutto ciò che si è detto per gli sviluppi di Fourier astratti si applica in questo caso particolare. Per esempio, vale l'IDENTITÀ DI PARSEVAL,

$$||f||_H^2 = \sum_{n=0}^{+\infty} |\langle f, p_n \rangle_H|^2\,.$$

Consideriamo ora questo problema: abbiamo costruito un sistema ortonormale massimale di polinomi in $L^2(a,b;\rho)$. E' ovvio che questo sistema non sia unico: moltiplicando tali polinomi per costanti di modulo 1, si trova ancora un sistema ortonormale massimale in $L^2(a,b;\rho)$. Mostriamo che però questo è l'unico modo per costruire un altro sistema $\{q_n(x)\}$ di *polinomi* ortonormali, massimale in $L^2(a,b;\rho)$: Notiamo prima di tutto che se $\{q_n(x)\}$ deve essere completo, per ogni $n \geq 0$, intero, uno almeno dei polinomi deve avere grado n.

Teorema 244 *Sia $\{q_n(x)\}$ un secondo sistema di polinomi ortonormali e massimale in $L^2(a,b;\rho)$. Sia $q_k(x)$ uno di tali polinomi, e questo abbia grado k. Esiste una costante α_k di modulo 1 tale che*

$$q_k(x) = \alpha_k p_k(x)\,.$$

<u>Dim.</u> Rappresentiamo

$$q_k(x) = \sum_{i=0}^{k} \alpha_i p_i(x)\,, \qquad \alpha_i = \int_a^b \rho(x) q(x) \overline{p_i}(x)\,\mathrm{d}x\,. \tag{3.30}$$

Si sa che per ogni $m < i$ si ha

$$\int_a^b x^m \rho(x) \overline{p_i}(x)\,\mathrm{d}x = 0\,.$$

Si fissi ora un indice i_0. Se $m < i_0$ allora $q_m \perp q_{i_0}$. Se $k < i_0$ allora

$$x^k \in \operatorname{span}\{q_0,\,\ldots\,,q_k\} \quad \text{perché } q_k \text{ ha grado } k$$

e quindi $x^k \perp q_{i_0}$. Dunque anche $p_k \perp q_{i_0}$. Da (3.30),

$$\alpha_{i_0} = \int_a^b \rho(x) q_k(x) \overline{p_{i_0}}(x)\,\mathrm{d}x = 0$$

per ogni $i_0 < k$; e quindi,

$$q_k(x) = \alpha_k p_k(x)\,.$$

Il modulo di α_k è 1 perché sia $p_k(x)$ che $q_k(x)$ hanno norma 1. ∎

Questo teorema prova che, assegnato $\rho(x)$, esiste sostanzialmente un solo sistema ortonormale massimale in $L^2(a,b;\rho)$. Di conseguenza, i polinomi $p_k(x)$ sono proporzionali a quelli che si ottengono ortogonalizzando, col metodo di Gram-Schmidt, i polinomi x^k. Il metodo di Gram-Schmidt applicato alla succesisone $\{x^k\}$ dà polinomi a coefficienti reali e quindi *non è restrittivo assumere che i polinomi $p_k(x)$ stessi abbiano coefficienti reali.*

I polinomi di Legendre e di Hermite sono definiti su intervalli simmetrici rispetto all'origine. Una conseguenza di ciò è la seguente:

Teorema 245 *I polinomi di Legendre ed Hermite di grado pari sono funzioni pari; quelli di grado dispari sono funzioni dispari.*

Dim. Come si è detto, nel caso dei polinomi di Legendre e di Hermite l'intervallo $(a, b) = (-a, a)$, $a = 1$ oppure $a = +\infty$, è un intervallo simmetrico rispetto a 0 e inoltre il peso è una funzione pari, si veda la tabella 3.1. Quindi, se $m < n$,

$$0 = \int_{-a}^{a} p_n(x)\rho(x)x^m \, \mathrm{d}x = (-1)^m \int_{-a}^{a} p_n(-x)\rho(x)x^m \, \mathrm{d}x$$

e quindi $p_n(-x)$ è ortogonale a tutti i polinomi di grado inferiore.

Il Teorema 244 mostra che $p_n(-x)$ è multiplo di $p_n(x)$:

$$p_n(-x) = C_n p_n(x) \, .$$

Identificando i coefficienti dei termini di grado più alto, si vede che $C_n = (-1)^n$.

Usiamo ora il principio di identità dei polinomi ed uguagliamo i coefficienti dei termini dello stesso grado a destra ed a sinistra. Si vede che per avere l'uguaglianza con n pari devono essere nulli i coefficienti degli addendi di grado dispari; se n è dispari devono essere nulli i coefficienti degli addendi di grado pari. ∎

3.4.3.2 Il calcolo dei polinomi ortogonali

Il calcolo dei polinomi ortogonali a partire dall'equazione differenziale è difficile. Per calcolarli esplicitamente conviene usare il metodi di Gram-Schmidt a partire dalla successione $\{x^n\}$ (con $n = 0\,,\, 1\,,\, \cdots$. Inoltre, è importante sapere che esiste una formula compatta per rappresentarli, a meno di una costante moltiplicativa:

Teorema 246 *Per ogni n, esiste una costante B_n tale che*

$$p_n(x) = B_n \frac{1}{\rho(x)} \frac{d^n}{dx^n} \left[\rho(x)\sigma^n(x)\right] \, . \tag{3.31}$$

Dim. Consideriamo una generica soluzione $y(x)$, non necessariamente polinomiale, dell'equazione (3.28), che riscriviamo:

$$\sigma(x)y'' + \tau(x)y' = \lambda y \tag{3.32}$$

Consideriamo la funzione $v_{(1)}(x) = y'(x)$. Questa verifica

$$\sigma(x)v''_{(1)} + \tau_{(1)}(x)v'_{(1)} = \lambda_{(1)}v_1$$

con

$$\tau_{(1)}(x) = \tau(x) + \sigma'(x)\,, \quad \lambda_{(1)}(x) = \lambda - \tau'(x)$$

(si noti che $\tau(x)$ ha grado 1 e quindi $\lambda_{(1)}$ è costante. Non si confonda $\lambda_{(1)}$ con λ_1 né i numeri $\lambda_{(k)}$ che ora introdurremo con λ_k).

In modo analogo, se $v_{(k)}(x) = y^{(k)}(x)$, si vede che $v_{(k)}(x)$ risolve

$$\sigma(x)v''_{(k)} + \tau_{(k)}(x)v_{(k)} = \lambda_{(k)}v_{(k)}\,, \tag{3.33}$$
$$\tau_{(k)}(x) = \tau(x) + k\sigma'(x)\,, \qquad \lambda_{(k)} = \lambda - k\tau'(x) - \frac{k(k-1)}{2}\sigma''(x)$$

(si noti che $\lambda_{(k)}$ è costante perché sono costanti sia $\tau'(x)$ che $\sigma''(x)$).

Si sa che, introducendo $\rho(x) > 0$ soluzione di

$$(\sigma(x)\rho)' = \tau(x)\rho\,, \tag{3.34}$$

la (3.32) si scrive nella forma

$$(\rho(x)\sigma(x)y')' = \lambda\rho(x)y\,.$$

Analogamente, sia $\rho_{(k)}(x)$ soluzione di

$$(\sigma(x)\rho_{(k)})' = \tau_{(k)}(x)\rho_{(k)}\,. \tag{3.35}$$

La $v_{(k)}(x)$ risolve

$$(\sigma(x)\rho_{(k)}(x)v'_{(k)})' = \lambda_{(k)}\rho_{(k)}v_{(k)}\,. \tag{3.36}$$

Mostriamo una relazione tra $\rho(x)$ e $\rho_{(k)}(x)$. Scrivendo esplicitamente le derivate del prodotto e dividendo rispettivamente per $\sigma(x)\rho(x)$ e $\sigma(x)\rho_{(k)}(x)$ i due membri di (3.34) e di (3.35) si vede che

$$\frac{\rho'(x)}{\rho(x)} = -\frac{\sigma'(x)}{\sigma(x)} + \frac{\tau(x)}{\sigma(x)}$$
$$\frac{\rho'_{(k)}(x)}{\rho_{(k)}(x)} = -\frac{\sigma'(x)}{\sigma(x)} + \frac{\tau_{(k)}(x)}{\sigma(x)} = -\frac{\sigma'(x)}{\sigma(x)} + \frac{\tau(x)}{\sigma(x)} + k\frac{\sigma'(x)}{\sigma(x)}$$

e quindi

$$\frac{\rho'_{(k)}(x)}{\rho_{(k)}(x)} = \frac{\rho'(x)}{\rho(x)} + k\frac{\sigma'(x)}{\sigma(x)}\,.$$

Ricordiamo che le funzioni $\rho(x)$ e $\rho_{(k)}(x)$, ambedue positive, sono definite a meno di un fattore costante positivo (si veda la tabella 3.1) e quindi, prendendo le primitive dei due membri e ponendo uguale a 0 la costante additiva, si trova

$$\rho_{(k)}(x) = \rho(x)\sigma^k(x)$$

e quindi anche

$$\rho_{(k)}(x) = \sigma(x)\rho_{(k-1)}(x)\,.$$

Combinando la (3.36) con queste relazioni, e posto $v_{(0)}(x) = y(x)$, $\rho_{(0)}(x) = \rho(x)$ e $\lambda_{(0)} = \lambda$, si vede che per ogni $k \geq 0$ vale

$$\lambda_{(k)}\rho_{(k)}(x)v_{(k)}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[\sigma(x)\rho_{(k)}(x)v'_{(k)}(x)\right] = \frac{\mathrm{d}}{\mathrm{d}x}\left[\rho_{(k+1)}(x)v'_{(k)}(x)\right]\,.$$

Ricordando che, per definizione, $v'_{(k)}(x) = v_{(k+1)}(x)$ si trova

$$\lambda_{(k)}\rho_{(k)}(x)v_{(k)}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[\rho_{(k+1)}(x)v_{(k+1)}(x)\right]\,.$$

Partendo ora da $k = 0$ e iterando si vede che

$$\rho(x)y(x) = \rho_{(0)}(x)v_{(0)}(x) = \frac{1}{\lambda}[\rho_{(1)}(x)v_{(1)}(x)]' = \frac{1}{\lambda}\left[\frac{1}{\lambda_{(1)}}\rho_{(2)}(x)v'_{(2)}(x)\right]'$$

e, iterando ulteriormente,

$$\begin{aligned}\rho(x)y(x) &= \frac{1}{\lambda_{(0)}\cdot\lambda_{(1)}\cdot\lambda_{(2)}\cdots\lambda_{(k-1)}}\frac{\mathrm{d}^k}{\mathrm{d}x^k}\left[\rho_{(k)}(x)v_{(k)}(x)\right]\\ &= \frac{1}{\lambda_{(0)}\cdot\lambda_{(1)}\cdot\lambda_{(2)}\cdots\lambda_{(k-1)}}\frac{\mathrm{d}^k}{\mathrm{d}x^k}\left[\rho(x)\sigma^k(x)v_{(k)}(x)\right]\,.\end{aligned}$$

Ripetiamo che questa relazione vale per ciascuna soluzione $y(x)$ della (3.28). Se però in particolare $y(x)$ è un polinomio di grado n, la sua n–ma derivata, ossia $v_{(n)}(x)$, è costante. Quindi, scrivendo la relazione precedente con $k = n$, si trova la (3.31). ∎

La formula (3.31) si chiama FORMULA DI RODRIGUEZ.

Usa scegliere reale la costante moltiplicativa così che, come si è già notato, *i polinomi ortogonali classici sono polinomi a coefficienti reali.*

La formula di Rodriguez dà un modo per il calcolo dei polinomi ortogonali, a meno di una costante di normalizzazione, che fa intervenire il calcolo di derivate. Se invece di richiedere ai polinomi di avere norma 1 richiediamo che siano "monici", ossia che il coefficiente del termine del grado maggiore sia 1, allora la costante nella formula di Rodriguez è uguale ad 1.

Il modo più efficiente per calcolare la successione dei polinomi ortogonali "monici" è di usare la formula di Rodriguez per calcolarne i primi, e per i successivi usare invece la relazione di ricorrenza che ora vediamo.

Il polinomio $xp_n(x)$ ha grado $n+1$, e quindi è combinazione lineare dei polinomi di grado minore od uguale a $n+1$:

$$xp_n(x) = \sum_{r=0}^{n+1} c_{r,n} p_r(x)\,. \tag{3.37}$$

Ricordiamo che un qualsiasi $p_j(x)$ è ortogonale a tutti i polinomi che hanno grado strettamente minore. Quindi,

$$c_{r,n} = \langle xp_n(x), p_r(x)\rangle = \langle p_n(x), xp_r(x)\rangle = 0 \qquad \text{se } r < n-1\,.$$

Quindi l'uguaglianza (3.37) si riduce a

$$c_{n+1,n}p_{n+1}(x) = xp_n(x) - c_{n,n}p_n(x) - c_{n-1,n}p_{n-1}(x)\,, \tag{3.38}$$

dove si intende $p_{-1}(x) = 0$ e anche $c_{-1,1} = 0$. D'altra parte, abbiamo detto di lavorare con polinomi monici. Confrontando i termini di grado più alto nei due membri, si vede che

$$c_{n+1,n} = 1$$

e quindi rimane

$$p_{n+1}(x) = xp_n(x) - c_{n,n}p_n(x) - c_{n-1,n}p_{n-1}(x)\,. \tag{3.39}$$

Mostriamo ora che i coefficienti $c_{n,n}$ e $c_{n-1,n}$ si possono calcolare se sono noti i polinomi di grado al più n. Basta per questo prendere il prodotto scalare dei due membri prima con $p_n(x)$ e poi con $p_{n-1}(x)$ per trovare

$$c_{n,n} = \frac{\langle xp_n(x), p_n(x)\rangle}{||p_n||^2}\,, \qquad c_{n-1,n} = \frac{\langle xp_n(x), p_{n-1}(x)\rangle}{||p_{n-1}||^2}\,.$$

Quindi, noti $p_0(x)$ e $p_1(x)$, polinomi "monici", si hanno tutte le informazioni necessarie per calcolare il polinomio monico $p_2(x)$; calcolato $p_2(x)$ si hanno tutte le informazioni per il calcolo di $p_3(x)$ ecc.

Noti i polinomi in forma "monica" e quindi anche le loro norme, si può calcolare la successione dei polinomi ortonormali.

Nella formula (3.39) compaiono due successioni, $a_n = c_{n,n}$ e $b_n = c_{n-1,n}$. Ricordando che si è posto $p_{-1}(x) = 0$ e $b_1 = c_{-1,1} = 0$, la (3.39) può riformularsi come segue:

Teorema 247 *Sia* $\{p_n\}$ *una successione di* polinomi monici *due a due ortogonali. Esistono due successioni* $\{a_n\}$ *e* $\{b_n\}$ *di numeri reali, con* $b_1 = 0$*, tali che*

$$xp_n(x) = p_{n+1}(x) + a_np_n(x) + b_np_{n-1}(x)\,.$$

In questa formula si intende $p_{-1}(x) = 0$.

Le due successioni possono calcolarsi in modo ricorsivo.

La formula (3.39) assume un aspetto particolarmente semplice nel caso dei *polinomi di Legendre* quando essi si scelgano in modo da avere $P_n(1) = 1$ (e quindi non necessariamente normalizzati e nemmeno monici). In questo caso sarà $c_{n+1,n} \neq 1$ Usiamo la (3.38) che vale anche se i polinomi non sono monici e ricordiamo che i polinomi di Legendre di grado pari sono funzioni pari e quelli di grado dispari sono funzioni dispari. E quindi

$$c_{n,n} = 0$$

ossia la (3.38) si scrive ora[8]:

$$\alpha_n P_{n+1}(x) = x P_n(x) - \beta_n P_{n-1}(x) \,. \tag{3.40}$$

Ricordiamo inoltre che $P_n(x)$ verifica

$$(1 - x^2) P_n''(x) - 2x P_n'(x) = -n(n+1) P_n(x)$$

e quindi, posto $x = 1$,

$$2P_n'(1) = n(n+1) P_n(1) = n(n+1) \,.$$

perché abbiamo deciso di imporre la condizione imporre la condizione $P_n(1) = 1$ per ogni n.

Poniamo ora $x = 1$ in (3.40) e nella sua derivata. Usando l'identità precedente si trova

$$\alpha_n + \beta_n = 1 \,, \qquad \alpha_n \frac{(n+1)(n+2)}{2} + \beta_n \frac{n(n-1)}{2} = 1 + \frac{n(n+1)}{2}$$

e quindi

$$\alpha_n = 1 - \frac{n}{2n+1} = \frac{n+1}{2n+1} \,, \qquad \beta_n = \frac{n}{2n+1} \,.$$

Si trova così la formula ricorsiva seguente, valida per i soli polinomi che risolvono l'equazione di Legendre e che verificano $P_n(1) = 1$:

$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}(x) \,.$$

[8]Ricordiamo dalla tabella 3.1 che la "P" maiuscola è la notazione standard per i polinomi di Legendre.

3.4.3.3 Gli zeri dei polinomi ortogonali classici

Gli zeri dei polinomi ortogonali classici hanno varie proprietà interessanti. In particolare:

Teorema 248 *Gli zeri dei polinomi ortogonali classici sono tutti reali, semplici ed appartengono all'intervallo* (a, b).

Dim. Gli zeri sono semplici perché una soluzione di un'equazione differenziale del secondo ordine che ha uno zero multiplo è identicamente nulla.

Si consideri il polinomio $p_n(x)$ e siano x_i, $0 \leq i \leq k$ i suoi zeri in (a, b). Sia, per assurdo, $k < n$.

Essendo gli zeri tutti semplici, il polinomio cambia segno quando si annulla. Sia quindi

$$q(x) = \Pi_{i=0}^{k}(x - x_i) \,.$$

Il polinomio

$$p(x)q(x)$$

ha segno costante.

D'altra parte, se $k < n$, $p_n(x)$ è ortogonale a $q(x)$ e quindi

$$0 = \int_a^b p_n(x)q(x)\rho(x)\,\mathrm{d}x\,,$$

in contrasto con la positività dell'integrando. ∎

Consideriamo ora le relazioni tra gli zeri di $p_n(x)$ e quelle di $p_k(x)$, con $k > n$.

Teorema 249 *Siano* α *e* β *due zeri consecutivi di* $p_n(x)$. *Il polinomio* $p_k(x)$ *si annulla nell'intervallo aperto* (α, β). *Inoltre,* $p_k(\alpha) \neq 0$ *e* $p_k(\beta) \neq 0$.

Dim. Per ipotesi, il polinomio $p_n(x)$ non si annulla, e quindi ha segno costante, in (α, β). Eventualmente cambiandone il segno, si può supporre che esso sia positivo.

Inoltre, $p_n(x)$ cambia segno in α ed in β perché i suoi zeri sono semplici. Di conseguenza,

$$p_n'(\alpha) > 0\,, \qquad p_n'(\beta) < 0\,.$$

Sia $k > n$ così che

$$\lambda_k < \lambda_n$$

(si ricordi la negatività degli autovalori). Supponiamo per assurdo che $p_k(x)$ non si annulli in (α, β). Se necessario si può cambiare il segno di $p_k(x)$ e supporre $p_k(x) > 0$ in (α, β).

Ricordiamo le uguaglianze

$$(\sigma(x)\rho(x)p_n')' = \lambda_n\rho(x)p_n\,, \qquad (\sigma(x)\rho(x)p_k')' = \lambda_k\rho(x)p_k$$

e introduciamo la funzione

$$W(x) = p_n(x)[\sigma(x)\rho(x)p_k'(x)] - [\sigma(x)\rho(x)p_n'(x)]p_k(x)\,.$$

Calcolandone la derivata si trova

$$W'(x) = [\lambda_k - \lambda_n]p_k(x)p_n(x)\rho(x) < 0\,.$$

Dunque,

$$W(\beta) - W(\alpha) < 0\,. \tag{3.41}$$

D'altra parte $p_n(\alpha) = p_n(\beta) = 0$ e quindi, direttamente dalla definizione di $W(x)$,

$$W(\beta) - W(\alpha) = -\left[\sigma(\beta)\rho(\beta)p_n'(\beta)\right]p_k(\beta) + \left[\sigma(\alpha)\rho(\alpha)p_n'(\alpha)\right]p_k(\alpha)\,. \tag{3.42}$$

Per quanto si è detto sopra sui segni dei polinomi e delle derivate di $p_n(x)$, si ha la somma di due addendi non negativi. La contraddizione trovata mostra che $p_k(x)$ deve avere almeno uno zero nell'intervallo (α, β).

Infine notiamo che la disuguaglianza (3.42) è stretta e quindi sia $p_k(\alpha)$ che $p_k(\beta)$ sono non nulli. ∎

La tabella 3.2 completa la tabella 3.1, introducendo due righe che riportano le equazioni differenziali e le formule di Rodriguez relativa alle tre classi di polinomi ortogonali.

3.4.4 Sviluppi in serie di Fourier-Bessel

Senza ancora specificare lo spazio in cui si opera, chiamiamo OPERATORE DI BESSEL di indice n l'operatore

$$y(\cdot) \longrightarrow (B_n y)(x) = \frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\left[xy'(x)\right] - \frac{n^2}{x^2}y(x)\,.$$

In certe applicazioni l'indice n è un qualsiasi numero complesso. Noi però ci limiteremo a considerare il caso degli indici interi; e dato che n figura elevato al quadrato, considereremo il caso degli indici n che sono *numeri interi non negativi.*

Tabella 3.2: Polinomi ortogonali classici ed equazioni differenziali

	Legendre: $P_n(x)$	Laguerre: $L_n(x)$	Hermite: $H_n(x)$
$\sigma(x)$	$(1-x^2)$	x	1
$\tau(x)$	$-2x$	$-x+1$	$-2x$
$\rho(x)$	1	e^{-x}	e^{-x^2}
λ_n	$-n(n+1)$	$-n$	$-2n$
Eq. Diff.	$(1-x^2)y''-2xy' = -n(n+1)y$	$xy''+(1-x)y' = -ny$	$y''-2xy' = -2ny$
Rodriguez	$\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n$	$e^x\frac{\mathrm{d}^n}{\mathrm{d}x^n}[e^{-x}x^n]$	$e^{x^2}\frac{\mathrm{d}^n}{\mathrm{d}x^n}[e^{-2x}]$
H	$L^2(-1,1)$	$L^2(0,+\infty;\rho)$	$L^2(-\infty,+\infty;\rho)$

Siano $y(x)$ ed $z(x)$ due autofunzioni corrispondenti ad autovalori $-\lambda$ e $-\mu$ tra loro diversi (il segno meno è introdotto solamente per conservare la notazione più comune in letteratura). Procediamo *in modo formale* a verificarne una proprietà di ortogonalità. Diciamo che stiamo operando "in modo formale" perché ancora non abbiamo specificato lo spazio in cui si intende lavorare e nemmeno il dominio dell'operatore B_n. D'altra parte, le equazioni per gli autovalori sono

$$\begin{array}{l} \frac{\mathrm{d}}{\mathrm{d}x}[xy'(x)] - \frac{n^2}{x}y(x) = -\lambda xy(x) \\ \frac{\mathrm{d}}{\mathrm{d}x}[xz'(x)] - \frac{n^2}{x}z(x) = -\mu xz(x)\,. \end{array} \tag{3.43}$$

Si osservi che l'equazione per l'autovalore λ può anche scriversi come

$$x^2y'' + xy' - (n^2 - \lambda x^2)y = 0\,. \tag{3.44}$$

Le equazioni del secondo ordine (3.43) hanno certamente soluzioni in ogni intervallo che non contiene 0. Dato che i coefficienti *non sono regolari in zero,* niente possiamo dire del comportamento delle soluzioni per x tendente a zero. Ciò nonostante, moltiplichiamo la prima per $\bar{z}(x)$ e moltiplichiamo per $y(x)$ la coniugata della seconda. Integriamo si $[0,1]$ e sottraiamo gli integrali. Notiamo che anche questo è un passaggio formale, che poi andrà giustificato, perché non sappiamo se gli integrali sono finiti.

Integrando per parti si trova l'uguaglianza

$$(\lambda - \bar{\mu}) \int_0^1 x y(x) \bar{z}(x) \, \mathrm{d}x = [x y(x) \bar{z}'(x) - y'(x) x \bar{z}(x)]_0^1 \, . \tag{3.45}$$

Un procedimento analogo è nella dimostrazione del Teorema 243. Però è opportuno vedere esplicitamente il calcolo per sottolineare il ruolo di n. Consideriamo il prodotto con $\bar{z}(x)$ della prima equazione in (3.43). L'integrale del prodotto si tratta come segue:

$$\begin{aligned} 0 &= \int_0^1 \left\{ \bar{z}(x) \frac{\mathrm{d}}{\mathrm{d}x}[x y'(x)] - \frac{n^2}{x} \bar{z}(x) y(x) + \lambda x \bar{z}(x) y(x) \right\} \mathrm{d}x \\ &= [x \bar{z}(x) y'(x)]_0^1 - \int_0^1 x \bar{z}'(x) y'(x) \, \mathrm{d}x - n^2 \int_0^1 \frac{1}{x} \bar{z}(x) y(x) \, \mathrm{d}x \\ &+ \lambda \int_0^1 x \bar{z}(x) y(x) \, \mathrm{d}x \, . \end{aligned}$$

Si vede da qui una differenza tra il caso $n = 0$ ed il caso $n > 0$: se $n > 0$, per dare senso a quest'uguaglianza, bisogna esplicitamente richiedere l'integrabilità di $\bar{z}(x) y(x)/x$.

Le funzioni $y(x)$ e $z(x)$ sono regolari per $x = 1$ ma potrebbero non esserlo per $x = 0$. L'uguaglianza precedente però mostra che, se i passaggi formali possono giustificarsi; se $[x \bar{z}(x) y'(x)]_0^1 = 0$ e se $\bar{\mu} \neq \lambda$, allora $y(x)$ e $z(x)$ sono ortogonali in $L^2(0, 1; \rho)$, con $\rho(x) = x$.

I passaggi precedenti sono giustificati se si richiede che le "autofunzioni" $y(x)$ e $z(x)$ appartengano al s.spazio delle funzioni $\phi(x) \in L^2(0, 1; \rho)$, $\rho(x) = x$, che verificano le proprietà seguenti:

$$\left\{ \begin{array}{l} \phi(x) \in W^{2,2}(\epsilon, 1) \quad \text{per ogni } \epsilon > 0 \, , \phi(1) = 0 \, , \\ \limsup_{x \to 0+} |\phi(x)| < +\infty \, , \qquad \limsup_{x \to 0+} |\phi'(x)| < +\infty \, , \\ B_n \phi \in L^2(0, 1; \rho) \, , \\ \frac{1}{\sqrt{x}} \phi(x) \in L^2(0, 1) \, , \quad \text{proprietà da richiedere solo se } n > 0. \end{array} \right. \tag{3.46}$$

Queste considerazioni suggeriscono di studiare gli operatori B_n nello spazio $X = L^2(0, 1; \rho)$, $\rho(x) = x$, attribuendo loro come dominio le funzioni che verificano le proprietà in (3.46).

Si ricordi che se una funzione *di una variabile* ha derivata (in senso debole) che è integrabile, essa è continua e quindi ha senso richiedere che il suo valore in 1 sia nullo. Niente possiamo dire per ora del valore in 0 perché la condizione $\phi''(x) \in L^2(\epsilon, 1)$ per ogni $\epsilon > 0$ non implica l'integrabilità di $\phi''(x)$ in un intorno di 0.

Andiamo ora a provare l'esistenza degli autovalori e delle autofunzioni, ed a giustificare i passaggi formali visti sopra. Conviene per questo introdurre l'equazione seguente, che si chiama EQUAZIONE DI BESSEL DI ORDINE n:

$$\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\left[xy'(x)\right]-\left(\frac{n^2}{x^2}-1\right)y(x)=0 \qquad \text{ossia} \qquad x^2y''+xy'-(n^2-x^2)y=0\,. \tag{3.47}$$

Questa è l'equazione degli autovalori scritta con $\lambda = 1$, ma *non* stiamo affermando che $\lambda = 1$ debba essere un autovalore. E' però vero che la conoscenza delle soluzioni dell'equazione di Bessel e dei loro zeri conduce ad identificare tutti gli autovalori e le corrispondenti autofunzioni di B_n. Mostriamo ciò nel caso $n > 0$ lasciando al lettore le considerazioni analoghe nel caso $n = 0$.

Supponiamo di poter trovare una soluzione $y(x)$ di (3.47), che verifica

$$\begin{cases} \lim_{x\to 0+} x|y(x)| = 0\,, \qquad \lim_{x\to 0+} |y'(x)| < +\infty\,, \\ \frac{1}{\sqrt{x}}y(x) \in L^2(0,1)\,. \end{cases}$$

Supponiamo inoltre che questa soluzione si annulli in un punto $c > 0$. In tal caso, è immediato trovare una autofunzione di B_n. Per vederlo poniamo $\phi(x) = y(cx)$. E' chiaro che $\phi(x)$ appartiena a $\operatorname{dom} B_n$. Inoltre, scrivendo l'uguaglianza (3.47) nel punto cx si trova

$$c^2x^2y''(cx)+cxy'(cx)-(n^2-c^2x^2)y(cx)=0\,.$$

Essendo

$$\phi'(x)=cy'(cx)\,,\quad \phi''(x)=c^2y''(cx)\,,$$

si vede che $\phi(x)$ risolve l'equazione (3.44) degli autovalori, con $\lambda = c^2$.

Vale anche il viceversa: sia $\phi(x)$ un'autofunzione che corrisponde ad un autovalore $c > 0$. Si noti che la $\phi(x)$ è certamente prolungabile a $(0, +\infty)$ perché $x = 0$ è l'unico punto nell'intorno del quale l'equazione non si può scrivere in forma normale. Posto $y_0(x) = \phi(x/c)$ si trova una soluzione di (3.47), con le proprietà dette sopra; e quindi possiamo concludere che *lo studio delle soluzioni dell'equazione di Bessel (3.47), e dei loro zeri, conduce ad identificare tutte le autofunzioni di* B_n.

Va osservata però una difficoltà che è questa: due autofunzioni $\phi(x)$ e $\psi(x)$ del medesimo autovalore c potrebbero condurre a soluzioni $y_0(x;\phi)$ e $y_0(x;\psi)$ di (3.47) del tutto indipendenti l'una dall'altra. Mostriamo che non è così e che $y_0(x;\phi)$ e $y_0(x;\psi)$ sono *tra loro proporzionali.* Ciò è conseguenza del teorema seguente.

Teorema 250 *Le soluzioni dell'equazione di Bessel di ordine n che sono limitate per $x \to 0+$ sono tutte e sole i multipli della funzione*

$$J_n(x) = \sum_{k=0}^{+\infty} (-1)^k \frac{1}{k!(n+k)!} \left(\frac{x}{2}\right)^{2k+n} .$$

Le funzioni $J_n(x)$ sono di classe $C^\infty(\mathbb{R})$ e inoltre:

$$\begin{cases} J_n(0) = 1 & \text{se } n = 0 \\ J_n(0) = 0\,, \quad J_n'(0) \neq 0 & \text{se } n > 0\,. \end{cases}$$

La dimostrazione si trova nel paragrafo 3.4.6.

E' immediata conseguenza di questo:

Corollario 251 *Se $\phi(x)$ e $\psi(x)$ sono autofunzioni di B_n col medesimo autovalore, allora $y_0(x;\phi)$ ed $y_0(x;\psi)$ sono tra loro proporzionali.*

E' così visto che lo studio degli autovalori ed autofunzioni dell'operatore B_n si riduce allo studio degli zeri della funzione $J_n(x)$.

La funzione $J_n(x)$ si chiama FUNZIONE DI BESSEL di ordine n (e di "prima specie". Le funzioni di Bessel "di seconda specie", che non tratteremo, sono particolari soluzioni illimitate dell'equazione di Bessel. Va notato però che esse verificano $\limsup_{x\to 0} \sqrt{x}|y(x)| = +\infty$ e quindi non possono usarsi per giustificare i calcoli che conducono alla (3.45)).

Per quello che sappiamo fino ad ora, l'operatore B_n potrebbe non avere autovalori, od averne un numero finito. Mostriamo invece l'esistenza di una successione di autovalori mostrando l'esistenza di una successione di zeri di ciascuna delle funzioni di Bessel. Premettiamo un'osservazione. Sia $y(x)$ una qualsiasi soluzione dell'equazione di Bessel di ordine n e sia

$$u(x) = \sqrt{x}y(x) \qquad \text{ossia} \qquad y(x) = \frac{1}{\sqrt{x}}u(x)\,.$$

E' chiaro che $y(x)$ ed $u(x)$ hanno i medesimi zeri positivi. E':

$$y'(x) = -\frac{1}{2}x^{-3/2}u(x) + x^{-1/2}u'(x)$$
$$y''(x) = \frac{3}{4}x^{-5/2}u(x) - x^{-3/2}u'(x) + x^{-1/2}u''(x)\,.$$

Sostituendo nell'equazione di Bessel si vede che $u(x)$ soddisfa l'equazione

$$u'' + q(x)u = 0\,, \qquad q(x) = \left[1 + \frac{1-4n^2}{4x^2}\right]\,. \tag{3.48}$$

Mostriamo:

Teorema 252 *Ogni soluzione $u(x)$ di (3.48), e quindi ogni soluzione $y(x)$ di (3.47), ha una successione (c_k) di zeri, $c_k \to +\infty$.*

Dim. Essendo $u(x)$ soluzione di un'equazione differenziale del secondo ordine di cui $x = 0$ è l'unico punto "singolare", gli zeri di $u(x)$ o sono in numero finito, o sono un insieme infinito. Se l'insieme degli zeri è infinito, esso può essere superiormente illimitato e eventualmente avere 0 come unico punto di accumulazione. L'asserto del teorema richiede solamente di mostrare l'esistenza di una successione divergente di zeri (mostreremo alla fine di questa parte la ragione per cui anche le soluzioni dell'equazione di Bessel che sono illimitate per $x \to 0$ non hanno zeri che si accumulano a zero). La proprietà cruciale che usiamo per questo è

$$q(x) > \frac{1}{2} \qquad \text{per } x \text{ sufficientemente grande.}$$

Sia x_0 tale che $q(x) > 1/2$ per $x \geq x_0$.

Notiamo che una soluzione non nulla di (3.48) cambia segno in ciascuno dei suoi zeri $c > 0$. Infatti, se si annullasse senza cambiare segno dovrebbe essere $u(c) = 0$ e anche $u'(c) = 0$; dunque, per il Teorema di Cauchy sulle equazioni differenziali ordinarie, si avrebbe $u(x) = 0$ per ogni x.

Supponiamo per assurdo che sia $u(x)$ una soluzione non nulla *che non cambia segno per x sufficientemente grande,* diciamo per $x \geq x_0$. Per fissare le idee sia $u(x) \geq 0$ per $x > x_0$. Allora, per $x \geq x_0$, si ha $u''(x) \leq 0$ e quindi $u'(x)$ decresce. Vogliamo provare che da qui segue la contraddizione seguente: la funzione $u(x)$ prende valori negativi. Per questo, introduciamo la funzione

$$v(x) = \frac{u'(x)}{u(x)}\,.$$

Questa funzione è ben definita per $x \geq x_0$ perchè abbiamo supposto che $u(x)$ non si annulli. Si vede immediatamente che verifica

$$v'(x) = -q(x) - v^2(x) \leq -q(x)\,.$$

Dunque,

$$v(x) \leq v(x_0) - \int_{x_0}^{x} q(s)\,\mathrm{d}s \leq v(x_0) - \frac{1}{2}(x - x_0) \to -\infty \quad \text{per} \quad x \to +\infty\,.$$

Dunque, $v(x)$ è negativa per x abbastanza grande, diciamo per $x \geq x_1$, *perché si è supposta la positività di $u(x)$.* E quindi $u'(x) < 0$ per $x > x_1$; e ricordiamo che $u'(x)$ *decresce.*

Dunque si trova la contraddizione

$$u(x) = u(2x_1) + \int_{2x_1}^{x} u'(s)\,\mathrm{d}s \le u(2x_1) + (x - 2x_1)u'(2x_1) \to -\infty \quad \text{per} \quad x \to +\infty\,.$$

Ciò prova che la $u(x)$ deve annullarsi in un punto $\tilde{x}_0 > x_0$ e quindi prendere valori negativi a destra di $\tilde{x}_0$.

Un argomento analogo al precedente mostra però che non può rimanere permanentemente negativa; ossia che $\tilde{x}_0$ non è l'ultimo degli zeri di $u(x)$. ∎

Si osservi che il teorema precedente si applica *a ciascuna soluzione dell'equazione di Bessel;* anche alle soluzioni che non sono regolari per $x \to 0+$. Le funzioni $J_n(x)$ essendo analitiche in $x = 0$ non hanno una successione di zeri che tende a zero. Ciò vale anche per le soluzioni che non sono regolari in 0, anche se la dimostrazione precedente non permette di dedurlo.

Quindi

Corollario 253 *Esiste una successione (λ_k, ϕ_k) di autovalori e autofunzioni dell'operatore B_n, con λ_k reale e $\phi_k(x)$ a valori reali. Inoltre, ϕ_k è ortogonale a ϕ_r nello spazio $L^2(0,1;\rho)$, $\rho(x) = x$, se $k \neq r$.*

Le funzioni $\phi_k(x)$ sono date da

$$\phi_k(x) = J_n(c_k x) \qquad \text{essendo } (c_k) \text{ la successione degli zeri di } J_n(x).$$

Si noti che le funzioni $\phi_k(x)$ così costruite *non* hanno norma unitaria in $X = L^2(0,1;\rho)$, $\rho(x) = x$; essendo però ortogonali due a due, ogni $f \in X$ appartenente allo spazio lineare che esse generano potrà scriversi nella forma

$$f(x) = \sum_{k=0}^{+\infty} f_k \phi_k(x) = \sum_{k=0}^{+\infty} f_k J_n(c_k x)\,, \qquad f_k = \frac{1}{||\phi_k||^2}\langle f, \phi_k\rangle \tag{3.49}$$

(norma e prodotto interno essendo quelli di X).

Serie di questo tipo si chiamano SERIE DI FOURIER-BESSEL.

E' importante conoscere il seguente risultato, che non proviamo:

Teorema 254 *Le autofunzioni di B_n sono un sistema completo in $L^2(0,1;\rho)$, con $\rho(x) = x$; e quindi ogni funzione $f(x)$ in tale spazio ammette uno sviluppo in serie di Fourier-Bessel.*

Da questo risultato segue l'importanza di conoscere le funzioni di Bessel ed i loro zeri, che sono tabulati con grande precisione.

Osservazione 255 Vale la pena di notare che anche gli sviluppi di Fourier di soli seni, o di soli coseni, si possono ottenere mediante una via analoga a quella seguita per gli sviluppi di Fourier-Bessel. Sia infatti $f(x)$ una funzione a quadrato integrabile su $(0,1)$. Come è noto, estendendola in modo dispari all'intervallo $(-1,1)$, si prova che essa è somma della serie di Fourier di soli seni

$$f(x)=\sum_{k=1}^{+\infty} f_k \sin k\pi x=\sum_{k=0}^{+\infty} f_k F(c_k x)$$

ove $F(x)=\sin x$ risolve

$$F''+F=0\,,\qquad F(0)=0$$

e i numeri $c_k=k\pi$ sono gli zeri non negativi di $F(x)$. Il primo di essi è nullo e ciò mostra che il primo termine della serie è nullo. ∎

3.4.5 Ulteriori proprietà delle funzioni di Bessel

Le funzioni di Bessel, e gli sviluppi di Fourier-Bessel, si incontrano nello studio dei fenomeni oscillatori, come le funzioni e^{ikx} e gli sviluppi di Fourier usuali; ma le funzioni e gli sviluppi di Fourier si incontrano per esempio nello studio delle vibrazioni di un filo mentre gli sviluppi di Fourier-Bessel si incontrano nello studio delle vibrazioni di una membrana. Visto che la membrana di un altoparlante riproduce con buona fedeltà il suono di un violino, possiamo chiederci quali similitudini ci siano tra le funzioni trigonometriche e le funzioni di Bessel. Una similitudine si è già vista: le funzioni di Bessel hanno infiniti zeri e inoltre sviluppo in serie di Fourier e di Fourier-Bessel hanno una struttura simile, si veda l'osservazione 255. Valgono inoltre le uguaglianze seguenti, che si ricavano immediatamente dalle serie che rappresentano le funzioni $J_n(x)$:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[x^{-n}J_n(x)\right]=-x^{-n}J_{n+1}(x)$$
$$\frac{\mathrm{d}}{\mathrm{d}x}\left[x^{n+1}J_{n+1}(x)\right]=x^{n+1}J_n(x)\,.$$

In particolare, se $n=0$, si trova

$$\frac{\mathrm{d}}{\mathrm{d}x}J_0(x)=-J_1(x)\,,\qquad \frac{\mathrm{d}}{\mathrm{d}x}xJ_1(x)=xJ_0(x)\,,$$

relazioni che "sostituiscono" quelle tra le derivate di $\sin x$ e $\cos x$. Una differenza importante è però questa:

Teorema 256 *Per ogni n esiste una costante M_n tale che*

$$|J_n(x)| < M_n \frac{1}{\sqrt{x}}\,.$$

Dim. ricordando l'equazione (3.48), basta provare che la funzione $u(x)$ rimane limitata.

Come si è detto, per $x > x_0$ è $q(x) > 1/2$. Moltiplichiamo per $u'(x)$ i due membri di (3.48). Si trova

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}x}[u'(x)]^2 = -q(x)\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}x}u^2(x) \leq -\frac{1}{4}\frac{\mathrm{d}}{\mathrm{d}x}u^2(x)\,. \tag{3.50}$$

Integrando i due membri si trova

$$[u'(x)]^2 + \frac{1}{2}u^2(x) = [u'(x_0)]^2 + \frac{1}{2}u^2(x_0)\,.$$

In particolare,

$$[u'(x)]^2 \leq [u'(x_0)]^2 + \frac{1}{2}u^2(x_0)\,, \qquad u^2(x) \leq 2[u'(x_0)]^2 + u^2(x_0)$$

come volevamo. ∎

Osservazione 257 Si noti che la dimostrazione precedente non usa la limitatezza per $x \to 0$ della soluzione e quindi vale per per ogni soluzione dell'equazione di Bessel; anche per quelle illimitate per $x \to 0+$. ∎

Infine, enunciamo senza dimostrazione il seguente risultato, che lega in modo più preciso le funzioni di Bessel con quelle trigonometriche: per ogni n esiste una funzione *limitata* $r_n(x)$ tale che

$$J_n(x) = \sqrt{\frac{2}{\pi x}}\cos\left(x - \frac{\pi}{4} - \frac{n\pi}{2}\right) + \frac{r_n(x)}{x^{3/2}}\,.$$

3.4.6 Le soluzioni dell'equazione di Bessel

E' immediato verificare che la serie di potenze proposta per $J_n(x)$ ha raggio di convergenza $+\infty$ e quindi può essere derivata termine a termine. Sostituendola nell'equazione di Bessel, si vede facilmente che $J_n(x)$ la risolve. E' però più interessante cercare di ricavarla. Semplifica lievemente i calcoli lavorare con

l'equazione (3.48). Ricerchiamo quindi una soluzione non identicamente nulla di tale equazione della forma

$$u(x)=\sum_{r=0}^{+\infty} u_r x^{r+m}$$

con m ed i coefficienti u_r da determinare. Imporremo la condizione $m \geq 0$ perché ricerchiamo soluzioni dell'equazione di Bessel che rimangono limitate per $x \to 0$.

E':

$$\begin{cases} 4x^2u(x) & = \sum_{r=2}^{+\infty} 4u_{r-2}x^{r+m} \\ 4x^2u''(x) & = 4m(m-1)u_0x^m+4m(m+1)u_1x^{m+1} \\ & \quad +\sum_{r=2}^{+\infty} 4(r+m)(r+m-1)u_r x^{r+m} \\ (1-4n^2)u(x) & = (1-4n^2)u_0x^m+(1-4n^2)u_1x^{m+1} \\ & \quad +\sum_{r=2}^{+\infty}(1-4n^2)u_r x^{r+m}\,. \end{cases}$$

Sostituendo nell'equazione, si trova una serie che si vuole essere nulla; ossia si vuole che i suoi coefficienti siano tutti nulli. Devono quindi valere le condizioni

$$\begin{aligned} &\left[4m(m-1)+1-4n^2\right]u_0=0 \\ &\left[4m(m+1)+1-4n^2\right]u_1=0 \\ &u_r=-\frac{4u_{r-2}}{4(r+m)(r+m-1)+1-4n^2}\,. \end{aligned}$$

Dunque, il valore di u_0 determina gli u_{2k} mentre il valore di u_1 determina i coefficienti u_{2k+1}. Se $u_0=u_1=0$ si trova la soluzione identicamente nulla, caso che non ci interessa. Quindi esaminiamo separatamente il caso $u_0 \neq 0$ ed $u_1=0$ dal caso $u_0=0$ ed $u_1 \neq 0$.

Se $u_1=0$ e $u_0=1$ allora m deve verificare

$$4m^2-4m+(1-4n^2)=0 \qquad \text{e quindi} \qquad m=n+\frac{1}{2}$$

(la soluzione $m=-n+1/2$ si esclude perchè si è chiesto $m \geq 0$).

In questo modo gli addendi della serie con $r=2k$ vengono ad essere

$$\begin{aligned} &u_{2k}x^{2k+n+1/2}\,, \\ &u_{2k}=-\frac{4u_{2(k-1)}}{4(2k+n+1/2)(2k+n-1/2)+1-4n^2} \\ &=-\frac{u_{2k-2}}{4k(k+n)}\,. \end{aligned} \tag{3.51}$$

Se invece si decide di lavorare con $u_0 = 0$ ed $u_1 = 1$ allora

$$m = n - \frac{1}{2}\,.$$

Questo caso quindi *non va considerato se* $n = 0$ perché conduce ad $m < 0$. Si può considerare se $n \geq 1$. In tal caso gli addendi della serie con $r = 2k+1$ vengono ad essere

$$u_{2k+1}x^{2k+n+1/2}\,, \qquad u_{2k+1} = -\frac{u_{2k-1}}{4(2k+n+1/2)(2k+n-1/2)+1-4n^2}\,.$$

Si vede quindi che i due casi apparentemente diversi conducono alla medesima soluzione.

Consideriamo quindi il caso $u_1 = 0$, $u_0 = 1$ ossia il caso (3.51). Si vede che

per $k = 1$ si ha $u_2 = -1/4(n+1)$
per $k = 2$ si ha $u_4 = +1/[4 \cdot 4 \cdot 2(n+2)(n+1)]$
e in generale $u_{2k} = (-1)^k(n-1)!/[4^k k!(n+k)!]\,.$

Dunque,

$$u(x) = 2^n[(n-1)!]\sqrt{x}\left[\sum_{k=0}^{+\infty}\frac{(-1)^k}{k!(n+k)!}\left(\frac{x}{2}\right)^{2k+n}\right]\,.$$

Dividendo per $2^n[(n-1)!]\sqrt{x}$ si trova la funzione $J_n(x)$.

L'equazione di Bessel, come l'equazione per $u(x)$, è un'equazione differenziale del secondo ordine e quindi ammette una seconda soluzione, linearmente indipendente dalla $J_n(x)$. Ovviamente, questa si ottiene dividendo per $\sqrt{x}$ una seconda soluzione della (3.50), linearmente indipendente dalla soluzione $u(x)$ appena trovata. Sembrerebbe naturale ricercare questa soluzione scegliendo per m il valore negativo che abbiamo scartato, $m = -n + 1/2$ e quindi ripetendo i calcoli precedenti. E' facile vedere che in questo modo *non* si trova una soluzione dell'equazione perché la relazione di ricorrenza che permette di costruire u_{2k} noto u_{2k-2} ha denominatore che si annulla per certi valori di k. Invece, nota la soluzione $u(x)$ di (3.50), ricerchiamo una seconda soluzione di forma

$$u(x)v(x)\,.$$

Sostituendo nell'equazione si vede che la funzione $v(x)$ deve soddisfare

$$u(x)v''(x) = -2u'(x)v'(x)\,.$$

Questa è un'equazione differenziale lineare del *primo ordine* per l'incognita $w(x) = v'(x)$. Quindi

$$w(x) = c\frac{1}{u^2(x)}\,.$$

Questa funzione non è limitata per $x \to 0+$ quando $n > 0$. Integrandola per esempio su un intervallo a cavallo di 1 si trova una funzione $v(x)$, definita per $x > 0$, anch'essa illimitata per $x \to 0+$ quando $n > 0$ perché l'integrale improprio di $1/u^2(x)$ non converge su un intervallo che ha 0 come estremo sinistro. Infatti, la $u(x)$ ha uno zero, per $x = 0$, di ordine $m = n + 1/2$. Moltiplicando l'integrale per $u(x)$ e quindi dividendo per $\sqrt{x}$ si ottiene una seconda soluzione dell'equazione di Bessel di ordine n, linearmente indipendente dalla $J_n(x)$ e si noti che questa è illimitata per $x \to 0+$ anche quando $n = 0$, e priva di zeri in un intorno di 0.

3.5 Il duale di uno spazio di Hilbert

Il paragrafo 2.7 riporta un elenco di spazi di Banach e ne rappresenta in modo "concreto" i duali. Può essere opportuno, ma non è indispensabile, conoscere queste rappresentazioni per la lettura di questo paragrafo.

Abbiamo già notato che, per ogni k fissato, il funzionale lineare

$$h \to \langle h, k \rangle$$

è continuo grazie alla *disuguaglianza di Schwarz*

$$|\langle h, k \rangle| \leq M||h|| \qquad \text{con} \qquad M = ||k|| \,.$$

Dunque la norma di questo funzionale non supera $||k||$ e *in realtà è uguale a* $||k||$ come si vede scegliendo

$$h = \frac{k}{||k||} \,.$$

Così come in dimensione finita, si mostra che questi funzionali esauriscono tutto il duale di H, ossia che H è un modello per il suo duale. Più precisamente vale:

Teorema 258 (di Riesz) *Sia ϕ un funzionale lineare e continuo su H. Esiste un* unico *$x_\phi \in H$ tale che*

$$\phi(h) = \langle h, x_\phi \rangle \qquad \forall h \in H \,. \tag{3.52}$$

La corrispondenza che a ϕ fa corrispondere x_ϕ è antilineare e inoltre

$$||\phi||_{H^*} = ||x_\phi||_H \,.$$

<u>Dim.</u> Si è appena detto che la trasformazione $h \to \langle h, y \rangle$ è lineare e continua su H, per ogni fissato $y \in H$. Ossia, almeno alcuni elementi del duale di H possono rappresentarsi come

$$\phi(h) = \langle h, y \rangle \,.$$

Mostriamo che questa rappresentazione, se esiste, è unica. Infatti sia

$$\phi(h) = \langle h, y \rangle = \langle h, x \rangle \qquad \forall h \in H \,.$$

Sottraendo, si trova $\langle h, x-y\rangle = 0$ per ogni $h \in H$ e quindi $x - y \perp H$, ossia $x - y = 0$.

Proviamo ora che *ogni* elemento ϕ di H^* si rappresenta come in (3.52).

Se $\phi = 0$ allora $x_\phi = 0$. Se $\phi \neq 0$,

$$\ker\phi \neq H$$

e la continuità di ϕ mostra che $\ker\phi$ è un s.spazio chiuso di H, diverso da H stesso. Dunque esiste $z \neq 0$, $z \perp \ker\phi$. Non è restrittivo assumere

$$||z|| = 1\,.$$

Si rappresenti ogni $h \in H$ nella forma[9]

$$h = \left(h - \frac{\phi(h)}{\phi(z)}z\right) + \frac{\phi(h)}{\phi(z)}z\,.$$

Essendo

$$\left(h - \frac{\phi(h)}{\phi(z)}z\right) \in \ker\phi\,, \qquad z \in [\ker\phi]^\perp$$

si ha

$$\langle h, [\overline{\phi(z)}z]\rangle = \langle \frac{\phi(h)}{\phi(z)}z, \overline{\phi(z)}z\rangle = \phi(h)\,.$$

Dunque,

$$x_\phi = \overline{\phi(z)}z\,.$$

Ciò prova che ogni $\phi \in H^*$ si rappresenta come in (3.52)

E' immediato verificare che la trasformazione $\phi \to x_\phi$, definita su H^*, è antilineare, usando l'unicità della rappresentazione. Infatti,

$$\begin{aligned}&(\alpha\phi + \beta\psi)(h) = \langle h, x_{\alpha\phi+\beta\psi}\rangle\,,\\ &(\alpha\phi + \beta\psi)(h) = \alpha\phi(h) + \beta\psi(h) = \alpha\langle h, x_\phi\rangle + \beta\langle h, x_\psi\rangle = \langle h, \overline{\alpha}x_\phi + \overline{\beta}x_\psi\rangle\,.\end{aligned}$$

L'unicità della rappresentazione implica

$$x_{\alpha\phi+\beta\psi} = \overline{\alpha}x_\phi + \overline{\beta}x_\psi\,.$$

Inoltre, si è notato che la norma della trasformazione $h \to \langle h, x_\phi\rangle$ è uguale a $||x_\phi||$. ∎

[9] questa rappresentazione mostra che $H = (\ker\phi) \oplus \operatorname{span}\{z\}$, fatto già notato nel Teorema 82.

Osservazione 259 E' importante notare che nella dimostrazione precedente il funzionale *continuo* ϕ potrebbe anche avere soltanto dominio denso in X. Anche in tal caso l'elemento x_ϕ può costruirsi, e il funzionale $h \to \langle h, x_\phi \rangle$ è *l'estensione per continuità di* ϕ *ad* H. Useremo quest'osservazione al teorema 264.

Notiamo inoltre che con le notazioni del paragrafo 2.9, la (3.52) si scrive

$$\phi(h) = \langle\langle \phi, h \rangle\rangle = \langle h, x_\phi \rangle . \quad \blacksquare$$

3.6 L'operatore aggiunto

In questa parte servono alcune nozioni del paragrafo 2.8. Ricordiamo che il grafico di un operatore A da H in K (generici spazi di Banach, ma in questa parte useremo spazi di Hilbert) è l'insieme

$$\mathcal{G}(A) = \{ (h, Ah) , \quad h \in \operatorname{dom} A \} \subseteq H \times K .$$

Operatori che hanno grafico chiuso si dicono OPERATORI CHIUSI. Useremo le seguenti proprietà: **1)** ogni operatore lineare continuo definito su un s.spazio chiuso è chiuso. **2)** un operatore lineare continuo e chiuso è definito su un s.spazio chiuso. **3)** Esistono operatori lineari continui definiti su tutto lo spazio la cui *immagine* non è chiusa. **4)** se A^{-1} esiste allora A è chiuso se e solo se A^{-1} è chiuso; **5)** se $A^{-1} \in \mathcal{L}(K, H)$ allora A è chiuso. **6)** un operatore chiuso con dominio uguale a tutto lo spazio è continuo.

Siano H e K due spazi di Hilbert e sia A un operatore lineare da H in K *anche NON continuo, ma con dominio denso in* H. Associamogli un operatore lineare da K in H che chiameremo OPERATORE AGGIUNTO. L'operatore aggiunto di A si indica col simbolo A^*.

Dobbiamo definire prima di tutto il *dominio* di A^*. Per definizione,

$$\operatorname{dom} A^* = \{ k \in K \mid \exists z \in H \text{ per cui } \langle Ah, k \rangle_K = \langle h, z \rangle_H \} .$$

Dunque, $k \in \operatorname{dom} A^*$ *se e solo se* $h \mapsto \langle Ah, k \rangle$ *ha estensione continua ad* H.

Vale:

Teorema 260 *L'elemento* z*, se esiste, è unico.*

Dim. Ne esistano due, z e ζ. Per ogni $h \in \mathrm{dom}\, A$ vale

$$\langle Ah, k\rangle_K = \langle h, z\rangle_H = \langle h, \zeta\rangle_H$$

e quindi

$$\langle h, z - \zeta\rangle_H \,.$$

Quest'uguaglianza vale *per ogni* $h \in \mathrm{dom} A$, che è *denso* in H. ciò implica che $\zeta = z$. ∎

E' quindi lecito definire

$$A^* k = z \,.$$

E' immediato verificare che *l'operatore* A^**, da* K *in* H*, è lineare.*

Vediamo *un'interpretazione geometrica* di questa definizione. Ricordiamo che

$$\begin{array}{l} k \in \mathrm{dom} A^* \iff \langle Ah, k\rangle_K - \langle h, z\rangle_H = 0 \qquad \forall h \in \mathrm{dom} A\,, \\ \text{e in tal caso} \quad z = A^* k\,. \end{array} \tag{3.53}$$

Queste proprietà si scrivono

$$\langle (h, Ah), (-z, k)\rangle_{H\times K} = 0 \qquad \forall (h, Ah) \in \mathcal{G}(A)\,.$$

Dunque $k \in \mathrm{dom} A^*$ e $z = A^* k$ se e solo se $(-z, k) \in [\mathcal{G}(A)]^\perp$. Ciò prova l'uguaglianza (3.54a) del prossimo teorema:

Teorema 261 *Valgono le uguaglianze seguenti:*

$$[\mathcal{G}(A)]^\perp = \{(-A^* k, k)\,,\ k \in \mathrm{dom} A^*\}\,. \tag{3.54a}$$

$$[\mathcal{G}(A^*)]^\perp = \mathrm{cl}\,\{(-Ah, h)\,,\ h \in \mathrm{dom} A\}\,. \tag{3.54b}$$

Dim. Rimane da provare la (3.54b). La (3.53) mostra che

$$\begin{array}{l} [\mathcal{G}(A^*)]^\perp \supseteq \{(-Ah, h)\,,\ h \in \mathrm{dom} A\} \\ \text{e quindi anche}\, [\mathcal{G}(A^*)]^\perp \supseteq \mathrm{cl}\,\{(-Ah, h)\,,\ h \in \mathrm{dom} A\}\,. \end{array}$$

Mostriamo che in realtà i due s.spazio sono uguali. Altrimenti esiste $(k_0, z_0) \neq (0, 0)$ tale che

$$(k_0, z_0) \in [\mathcal{G}(A^*)]^\perp \cap \{(-Ah, h)\,,\ h \in \mathrm{dom} A\}^\perp\,.$$

La condizione $(k_0, z_0) \in \{(-Ah, h)\,,\ h \in \mathrm{dom} A\}^\perp$ si scrive

$$\langle k_0, Ah\rangle_K = \langle z_0, h\rangle_H \text{ e quindi } k_0 \in \mathrm{dom} A^* \text{ e } z_0 = A^* k_0$$

ossia $(k_0, z_0) \in \mathcal{G}(A^*)$. Quindi si hanno le due inclusioni

$$(k_0, z_0) \in [\mathcal{G}(A^*)] \cap [\mathcal{G}(A^*)]^{\perp} \quad \text{così che } (k_0, z_0) = (0, 0)\,.$$

Ciò prova l'uguaglianza in (3.54b). ∎

Notiamo che

$$\{(-Ah, h)\,,\ h \in \mathrm{dom} A\} \text{ è chiuso se e solo se } \{(h, Ah)\,,\ h \in \mathrm{dom} A\} \text{ è chiuso}\,.$$

Questa osservazione e la (3.54b) danno:

Corollario 262 *Se A: $H \mapsto K$ ha dominio denso ed è chiuso allora*

$$[\mathcal{G}(A^*)]^{\perp} = \{(-Ah, h)\,,\ h \in \mathrm{dom} A\}\,.$$

Usiamo ora nuovamente il fatto che l'insieme ortogonale di un qualsiasi insieme è chiuso. La (3.54a) dice che $\{(-A^*k, k)\,,\ k \in \mathrm{dom} A^*\}$ è chiuso e quindi anche

$$\mathcal{G}(A^*) = \{(k, A^*k)\,,\ k \in \mathrm{dom} A^*\} \quad \text{è chiuso}$$

ossia:

Teorema 263 *Sia A: $H \mapsto K$ un operatore lineare con dominio denso in H. L'operatore A^* è chiuso.*

Sottolineiamo che A^* è chiuso anche se A non è chiuso.

Studiamo ora le proprietà degli operatori aggiunti considerando prima di tutto gli aggiunti degli operatori continui.

3.6.1 L'aggiunto di un operatore continuo

Notiamo prima di tutto:

Teorema 264 *Se A è lineare e* continuo *da H in K, con dominio* denso *in H, allora il suo aggiunto ha dominio* uguale *a K.*

<u>**Dim.**</u> Infatti, il funzionale

$$h \to \langle Ah, k\rangle$$

è continuo per ogni k e quindi, per il *teorema di Riesz* e l'osservazione 259, si rappresenta nella forma $\langle h, z\rangle$. ∎

Vale:

Lemma 265 *Sia A lineare e continuo da H in K, con dominio denso in H. Allora, $A^* \in \mathcal{L}(K,H)$ e $||A^*||_{\mathcal{L}(K,H)} \le ||A||$.*

<u>Dim</u>. Si è già notato che A^* è definito su K. Dalla disuguaglianza di Schwarz,

$$||A^*k|| = \sup_{||h||=1} \langle h, A^*k\rangle = \sup_{||h||=1} \langle Ah, k\rangle \le \sup_{||h||=1} ||Ah|| \cdot ||k|| = ||A|| \cdot ||k|| \,.$$

Dunque, A^* è un operatore limitato e

$$||A^*|| \le ||A|| \,. \quad ∎$$

Dato che dom$A^* = K$ possiamo calcolare $A^{**} = (A^*)^*$. Dal lemma precedente, $||A^{**}|| \le ||A^*||$. Proviamo ora:

Teorema 266 *Sia A lineare e continuo da H in K, con dominio denso in H. L'operatore A^{**} è l'estensione continua di A ad H e quindi, in particolare,*

$$||A|| = ||A^*|| \,.$$

<u>Dim</u>. Si sa già che A^{**} è definito su H. Proviamo che estende A. Per questo consideriamo il funzionale

$$k \to \langle A^*k, h\rangle_H \,.$$

Per definizione, se $h \in \operatorname{dom} A$, questo è uguale a

$$\langle k, Ah\rangle_H$$

e quindi $h \in \operatorname{dom} A^{**}$, con $A^{**}h = Ah$, ossia A^{**} estende A.

Di conseguenza vale anche $||A^*|| \le ||A|| \le ||A^{**}|| \le ||A^*||$ e quindi $||A|| = ||A^*||$. ∎

Osservazione 267 Il Teorema 264 ha come conseguenza immediata la continuità dell'operatore aggiunto perché un operatore chiuso definito su uno spazio di Banach è continuo per il Teorema 106. La dimostrazione che abbiamo data qui è elementare, nel senso che non dipende dal Teorema di Baire, ed inoltre identifica $\|A^*\|$. ∎

Notiamo ora che se $A \in \mathcal{L}(H,K)$ (e quindi con dominio H) vale

$$\langle Ah, k\rangle_K = \langle h, A^*k\rangle_H \qquad \forall h \in H \,,\ k \in K \,.$$

Usando questa uguaglianza è facile verificare che valgono le seguenti regole di calcolo:

Teorema 268 *Sia $A \in \mathcal{L}(H,K)$. Vale:*

$$(\alpha A)^* = \bar{\alpha} A^* \,; \qquad (A+B)^* = A^* + B^* \,.$$

Se A^{-1} esiste allora esiste anche $(A^*)^{-1}$ e vale

$$(A^{-1})^* = (A^*)^{-1} \,. \tag{3.55}$$

Sia $B \in \mathcal{L}(H,K)$ e $A \in \mathcal{L}(K,W)$. Si ha:

$$(AB)^* = B^* A^* \,.$$

Una forma più generale della (3.55) sarà provata nel Teorema 283. Le altre proprietà sono ovvie.

3.6.2 Il caso generale degli operatori illimitati

I risultati di questo paragrafo sono alla base di pressoché tutti i risultati seguenti. La lettura di questo paragrafo è delicata e si può omettere in una prima fase dello studio. Il Lemma 285 e il Teorema 286 del paragrafo 3.6.3 sommarizzano i risultati cruciali nella forma che verrà usata in seguito.

Quanto si è visto al paragrafo 3.6.1 ci dice che l'operatore aggiunto A^* di A: $H \mapsto K$ è continuo se e solo se A è continuo, e in tal caso dom$A^* = K$ anche se il dominio di A è solamente denso in H. Dunque se un operatore è illimitato anche il suo aggiunto è illimitato e per questa ragione lo studio degli operatori aggiunti è alquanto delicato. Però, ogni operatore aggiunto è chiuso e quindi aiutano le considerazioni geometriche che hanno condotto ai teoremi 261 e 263. Per questa ragione premettiamo alcune considerazioni sugli operatori chiusi.

3.6.2.1 Premesse sugli operatori chiusi

I risultati che presentiamo in questo paragrafo valgono per operatori tra generici spazi di Banach che indicheremo con i simboli H e K perché li useremo quando gli spazi sono di Hilbert.

Gli operatori lineari che si incontrano in questo paragrafo non hanno dominio denso. Quando questa ipotesi serve, essa è esplicitamente richiesta.

Prima di tutto notiamo che un operatore potrebbe non essere chiuso perché gli viene assegnato un dominio troppo piccolo. Per esempio l'operatore identità

(con dominio tutto lo spazio) è ovviamente chiuso ma se ne viene considerata la restrizione ad un sottospazio denso ma non chiuso l'operatore che si ottiene non è chiuso: è un operatore che ammette un'estensione chiusa.

Un operatore lineare che ammette una estensione che è un operatore chiuso si chiama CHIUDIBILE.

A questo proposito valgono i risultati seguenti, di ovvia dimostrazione:

Teorema 269 *Le proprietà seguenti valgono se H e K sono spazi di Banach; quindi in particolare anche se sono spazi di Hilbert:*

1. *un sottoinsieme $\mathcal{G}$ di $H \times K$ è grafico di un operatore lineare se e solo se è un s.spazio di $H \times K$ che non contiene elementi $(0, k)$ con $k \neq 0$;*

2. *Se A_0 estende A allora $\mathcal{G}(A_0) \supseteq \mathcal{G}(A)$. In particolare, se A_0 è estensione chiusa di A allora*
$$\mathcal{G}(A_0) \supseteq \mathrm{cl}\,[\mathcal{G}(A)]\ ;$$

3. *Un operatore lineare A da H in K ammette estensioni chiuse se e solo se $\mathrm{cl}\,\mathcal{G}(A)$ è grafico di un operatore, ossia se non contiene elementi della forma $(0, k)$ con $k \neq 0$. In tal caso l'operatore il cui grafico è $\mathrm{cl}\,\mathcal{G}(A)$ è la estensione chiusa di A che ha il* minimo *grafico*[10]*. Essa si chiama la* MINIMA ESTENSIONE CHIUSA DI A.

4. *la minima estensione chiusa A di A_0 può costruirsi in questo modo: si considerano le successioni $(\,(h_n, A_0 h_n)\,)$* che sono convergenti in $H \times K$*. Allora,* $\mathrm{dom} A$ *è il s.spazio di H i cui elementi sono i limiti delle successioni (h_n) e*
$$A\,(\lim h_n) = \lim A h_n\,.$$

Elenchiamo esplicitamente due conseguenze importanti.

Combinando le proprietà 1 e 3 si trova:

Corollario 270 *La minima estensione chiusa di un operatore lineare e chiudibile è essa stessa un operatore lineare.*

La proprietà 4 del Teorema 269 in particolare mostra che il dominio della minima estensione chiusa di A è contenuto in $\mathrm{cl}\,\mathrm{dom} A$. Ciò prova la prima affermazione del corollario seguente:

Corollario 271 *Sia A: $H \mapsto K$ è chiudibile e sia $\mathcal{A}$ la sua minima estensione chiusa. si ha:*

[10] minimo nel senso dell'ordine parziale dato dall'inclusione di insiemi.

1. *Il dominio della minima estensione chiusa* $\mathcal{A}$ *è denso in* H *se e solo se* $\mathrm{dom}A$ *è denso in* H*.*

2. *se* $\mathcal{A}$ *è invertibile anche* A *è invertibile e* $\mathcal{A}^{-1}$ *è la minima estensione chiusa di* A^{-1}*.*

<u>Dim</u>. Rimane da provare l'asserto 2. L'invertibilità di A è ovvia perché se $Ah_0 = 0$ si ha anche $\mathcal{A}h_0 = 0$ e quindi $h_0 = 0$.

Essendo $\mathcal{A}$ la minima estensione chiusa di A, $\mathcal{G}(\mathcal{A}) = \mathrm{cl}\,\mathcal{G}(A)$. D'altra parte

$$\mathcal{G}\left(\mathcal{A}^{-1}\right) = \{(\mathcal{A}h, h)\,,\ h \in \mathrm{dom}\mathcal{A}\}\,, \qquad \mathcal{G}\left(A^{-1}\right) = \{(Ah, h)\,,\ h \in \mathrm{dom}A\}$$

e quindi

$$\mathcal{G}\left(\mathcal{A}^{-1}\right) = \mathrm{cl}\,\mathcal{G}\left(A^{-1}\right)\,. \quad \blacksquare$$

E' ovvio che le estensioni chiuse di un operatore non sono necessariamente uniche: certamente non lo sono quando l'operatore è continuo ma il suo dominio non è denso. Mostriamo un fatto importante: *le estensioni chiuse, se esistono, non sono uniche nemmeno se il dominio è denso, ma diverso da tutto lo spazio.* Premettiamo un lemma:

Lemma 272 *Sia* H *uno spazio di Banach e sia* X_0 *un suo s.spazio chiuso. Per ogni* $\xi \in H$*, il sottospazio*

$$\mathrm{span}\,\{X_0, \xi\} = \{x + \alpha\xi\,, \quad x \in X_0\,,\ \alpha \in \mathbb{C}\}$$

è chiuso.

<u>Dim</u>. L'asserto è ovvio se $\xi \in X_0$ perché in tal caso $\mathrm{span}\,\{X_0, \xi\} = X_0$. Altrimenti, consideriamo una successione $(x_n + \alpha_n\xi)$ con $x_n \in X_0$ e $\alpha_n \in \mathbb{C}$. Se questa successione non converge niente dobbiamo fare. Se invece converge, dobbiamo provare che il suo limite appartiene a $\mathrm{span}\,\{X_0, \xi\}$. Esista quindi

$$y = \lim(x_n + \alpha_n\xi)\,.$$

Consideriamo due casi:

caso $\{\alpha_n\}$ **limitata** in questo caso passando ad una s.successione possiamo supporre $\alpha_n \to \alpha_0$ così che anche $x_n \to x_0$ ed $y = x_0 + \alpha_0\xi \in \mathrm{span}\,\{X_0, \xi\}$, come si voleva.

caso $\{\alpha_n\}$ illimitata In questo caso passando ad una s.successione si può supporre $|\alpha_n| \to +\infty$.

Indichiamo $y_n = x_n + \alpha_n \xi$. La successione (y_n) è limitata perché è convergente e quindi $\lim y_n/\alpha_n = 0$. Dunque,

$$\frac{x_n}{\alpha_n} = \frac{y_n}{\alpha_n} - \xi \to -\xi$$

e quindi ξ è limite di una successione in X_0, che è chiuso. Dunque $\xi \in X_0$ e span $\{X_0, \xi\} = X_0$, chiuso. ∎

Proviamo ora:

Teorema 273 *Questo risultato vale se H e K sono spazi di Banach e quindi in particolare anche se sono spazi di Hilbert. Sia A_0 un operatore* chiuso *da H in K il cui dominio non è uguale ad H. Esso ammette estensioni chiuse con dominio strettamente più grande.*

Dim. Si scelga un qualunque $h_0 \notin \mathrm{dom} A_0$ ed un qualunque $k_0 \in K$. Definiamo

$$\mathrm{dom} A = \{h + \alpha h_0\,, \quad h \in \mathrm{dom} A_0\,,\ \alpha \in \mathbb{C}\} \qquad A(h + \alpha h_0) = A_0 h + \alpha k_0\,.$$

Notiamo che la definizione è ben data perché $h + \alpha h_0 = h_1 + \beta h_0$ con h ed h_1 in $\mathrm{dom} A_0$ implica $h = h_1$ ed $\alpha = \beta$ perché $h_0 \notin \mathrm{dom} A_0$.

L'operatore A è lineare, estende A_0 ed ha dominio strettamente più grande di quello di A_0. Il suo grafico è

$$\begin{aligned}
&\mathcal{G}(A) = \{(h + \alpha h_0, A_0 h + \alpha k_0) \quad h \in \mathrm{dom}\, A_0 \ \ \alpha \in \mathbb{C}\} \\
&= \{(h, A_0 h) + \alpha(h_0, k_0) \quad h \in \mathrm{dom}\, A_0 \ \ \alpha \in \mathbb{C}\} \\
&= \mathcal{G}(A_0) + \{\alpha(h_0, k_0) \ \ \alpha \in \mathbb{C}\} = \mathrm{span}\ \{\mathcal{G}(A_0), (h_0, k_0)\}
\end{aligned}$$

chiuso per il Lemma 272. ∎

Si noti che l'asserto del Teorema 273 vale anche se l'operatore ha dominio denso.

In particolare, il Teorema 273 mostra che un operatore chiudibile con dominio diverso da tutto lo spazio ha infinite estensioni chiuse e che *il grafico di ogni estensione chiusa di A_0 contiene* $\mathrm{cl}\,\mathcal{G}(A_0)$. Dunque, come già notato nel Teorema 269:

Teorema 274 *La* minima *estensione chiusa di un operatore chiudibile A_0 è quella in cui grafico è* $\mathrm{cl}\,\mathcal{G}(A_0)$.

3.6.2.2 Aggiunti di operatori illimitati

Torniamo ora a considerare operatori lineari che operano tra spazi di Hilbert.

Sia $A\colon H \mapsto K$ un operatore lineare con dominio denso in H così che esiste l'operatore A^*. Ricordiamno, dal Teorema 263, che A^* è chiuso. Però potrebbe accadere che il dominio di A^* sia "molto piccolo" e in tal caso A^* dà poche informazioni. Vediamo un esempio che è un caso specifico di una proprietà generale che vedremo al Corollario 278.

Esempio 275 In quest'esempio conviene distinguere gli elementi $[x]$ di $L^2(0,1)$, che sono classi di equivalenza, dal rappresentante x della classe.

Sia $H = L^2(0,1)$ e sia

$$\mathrm{dom}\, A = \{[x] \in L^2(0,1) \text{ con rappresentante continuo}\}\,.$$

Sia x il rappresentante continuo di $[x] \in \mathrm{dom}\, A$. Definiamo

$$A[x] = x(0)\,.$$

Ossia, A è un *funzionale lineare* con dominio denso. Se $k \in \mathbb{C}$ è nel dominio di A^*, esiste $[z] \in L^2(0,1)$ per cui

$$\bar{k}h(0) = \int_0^1 \bar{z}(s)h(s)\,\mathrm{d}s \qquad \forall h \in L^2(0,1)\,.$$

Ciò può solo aversi se $k = 0$ (e allora anche $z = 0$); ossia, $\mathrm{dom}\, A^* = \{0\}$. ∎

Dunque l'operatore A^* può avere dominio "piccolo". Consideriamo invece il caso in cui A^* ha dominio denso in K. Allora si può definire A^{**}. Vale:

Teorema 276 *Se* $\mathrm{dom} A^*$ *è denso in* K *allora l'operatore* A *è chiudibile ed* A^{**} *è la* minima *estensione chiusa di* A.

Dim. Scriviamo la (3.54b) con A^* al posto di A e quindi A^{**} al posto di A^*.

$$[\mathcal{G}(A^{**})]^\perp = \mathrm{cl}\,\{(-A^*k, k)\,,\ k \in \mathrm{dom} A^*\} = [\mathcal{G}(A)]^\perp\,.$$

La seconda uguaglianza è la (3.54a).

Passando agli ortogonali si ha

$$\mathcal{G}(A^{**}) = \left[[\mathcal{G}(A^{**})]^\perp\right]^\perp = \left[[\mathcal{G}(A)]^\perp\right]^\perp = \mathrm{cl}\,\mathcal{G}(A)\,.$$

Dunque la chiusura del grafico di A è un grafico, ossia A è chiudibile, e la sua minima estensione chiusa è proprio A^{**}. ∎

Osservazione 277 Si noti quindi che se A^* è continuo con dominio denso anche A è continuo; e ciò spiega perché nel caso dell'esempio 275 il dominio dell'aggiunto deve essere 0: se il dominio fosse $\mathbb{R}$ l'operatore A^* sarebbe continuo e si potrebbe definire A^{**}, estensione *continua* di A, che invece *non* è continuo. ∎

Quest'argomento si può ripetere per ogni *funzionale* lineare:

Corollario 278 *L'aggiunto di un* funzionale lineare non continuo *ha dominio uguale a* 0*.*

Notiamo:

Teorema 279 *Sia A_0: $H \mapsto K$ un operatore chiudibile con dominio denso e sia A la sua minima estensione chiusa. Si ha:*

$$A_0^* = A^* .$$

Dim. Si ha:

$$[\mathrm{cl}\,\{(-Ah,h)\ \ h\in \mathrm{dom}A\}]^{\perp} = [\{(-Ah,h)\ \ h\in \mathrm{dom}A\}]^{\perp}$$
$$= [\{(-A_0h,h)\ \ h\in \mathrm{dom}A_0\}]^{\perp}$$

La (3.54b) mostra che A^* ed A_0^* hanno il medesimo grafico e quindi $A_0^* = A^*$. ∎

E ora proviamo:

Teorema 280 *Se A è* chiuso con dominio denso *anche A^* è* chiuso con dominio denso; *e quindi A^{**} può definirsi, ed è uguale ad A.*

Dim. In questa dimostrazione useremo l'uguaglianza:

$$[\mathcal{G}(A^*)]^{\perp} = \{(-Ah,h)\,,\ h\in \mathrm{dom}A\}$$

che vale perché l'operatore A è chiuso (si veda il Corollario 262).

Sia $k \in K$ ortogonale a dom A^*. In tal caso,

$$(k,0)\in [\mathcal{G}(A^*)]^{\perp} = \{((-Ah,h)\,,\ h\in \mathrm{dom}A\}$$

e quindi $k = 0$. Dunque, dom A^* è denso in K. ∎

I risultati appena visti mostrano che la classe degli peratori chiusi è la classe "naturale" in cui studiare l'operatore aggiunto.

3.6.2.3 Regole di calcolo per gli operatori aggiunti

Presentiamo ora alcune regole di calcolo per gli operatori aggiunti. Le dimostrazioni delle proprietà elencate nel Teorema 281 sono facili e sono posposte nel paragrafo 3.6.6 ma si consiglia di svolgerle per esercizio.

Teorema 281 *Siano A e B operatori lineari operanti tra gli spazi che specificheremo e ambedue con dominio denso così che i loro aggiunti A^* e B^* esistono. Si ha:*

a) *L'operatore A opera da H in K ed $\alpha \in \mathbb{C}$. Si ha:*

$$(\alpha A)^* = \bar{\alpha} A^* .$$

b) *Gli operatori A e B operano da H in K. Supponiamo che $\operatorname{dom}(A+B)$ sia denso in H così che $(A+B)^*$ esiste. Se $k \in (\operatorname{dom} A^*) \cap (\operatorname{dom} B^*)$ allora $k \in \operatorname{dom}(A+B)^*$ e*

$$(A+B)^* k = A^* k + B^* k ;$$

Ossia, $(A+B)^*$ estende $A^* + B^*$. In particolare, se i due operatori A e B sono continui, $(A+B)^* = A^* + B^*$.

c) *Sia B da H in K e A operi da K in W. Ricordiamo che, per la definizione di operatore composto, $\operatorname{dom}(AB) = \{h \in \operatorname{dom} B \ : \ Bh \in \operatorname{dom} A\}$. Supponiamo che $\operatorname{dom} AB$ sia denso in H così che $(AB)^*$ esiste. Allora*

Allora,

$$(AB)^* \text{ estende } B^* A^* .$$

Se $A \in \mathcal{L}(K, W)$ e $B \in \mathcal{L}(H, K)$ allora $(AB)^* = B^* A^*$.

I risultati seguenti, relativi agli inversi, sono particolarmente importanti e saranno richiamati al paragrafo 3.6.3.

Teorema 282 *Sia A: $H \mapsto K$ con dominio denso. Si ha:*

$$A^* k = 0 \iff k \in [\operatorname{im} A]^\perp .$$

<u>Dim.</u> Infatti, $k_0 \perp \mathcal{G}(A)$ significa

$$0 = \langle k_0, Ah \rangle_K \qquad \forall h \in \operatorname{dom} A .$$

Ciò equivale a $k_0 \in \operatorname{dom} A^*$ e $A^* k_0 = 0$. ∎

Teorema 283 *Sia A lineare da H in K con dominio denso e supponiamo che: 1)* $\ker A = 0$*; 2)* $\operatorname{im} A$ *denso in K. Allora A^* ammette inverso e vale*

$$(A^*)^{-1} = \left(A^{-1}\right)^* . \tag{3.56}$$

Dim. Ricordiamo che $\mathcal{G}(A^{-1}) = \{(Ah, h)\,,\ h \in \mathrm{dom} A\}$. Dunque, la (3.54b) si scrive

$$[\mathcal{G}(A^*)]^\perp = \mathrm{cl}\,\mathcal{G}(-A^{-1})\,. \tag{3.57a}$$

La (3.54a) è

$$[\mathcal{G}(A)]^\perp = \{(-A^*k, k)\,,\ k \in \mathrm{dom} A^*\} \tag{3.57b}$$

e se accade che l'insieme a destra è grafico di un operatore allora $-A^*$ è invertibile e $[\mathcal{G}(A)]^\perp = \mathcal{G}(-(A^*)^{-1})$. Per vedere che $\{(-A^*k, k)\,,\ k \in \mathrm{dom} A^*\}$ è un grafico proviamo:

$$(0, k_0) \in \{(-A^*k, k)\,,\ k \in \mathrm{dom} A^*\} = [\mathcal{G}(A)]^\perp \implies k_0 = 0\,.$$

Infatti, $(0, k_0) \perp \mathcal{G}(A)$ significa

$$0 = \langle 0, h\rangle_H + \langle k_0, Ah\rangle_K = \langle k_0, Ah\rangle_K$$

e ciò implica $k_0 = 0$ perché l'immagine di A è densa in K.

Dunque $-A^*$ e quindi anche A^* sono invertibili e si hanno ambedue le uguaglianze seguenti:

$$[\mathcal{G}(A^*)]^\perp = \mathrm{cl}\,\mathcal{G}(-A^{-1})\,, \tag{3.58a}$$

$$[\mathcal{G}(A)]^\perp = \mathcal{G}((-A^*)^{-1})\,. \tag{3.58b}$$

L'operatore $-A^{-1}$, che opera da K in H, ha le medesime proprietà di A e quindi queste uguaglianze valgono ambedue anche per $-A^{-1}$. In particolare, la (3.58b) scritta per $-A^{-1}$ è

$$[\mathcal{G}(-A^{-1})]^\perp = \mathcal{G}\left(((A^{-1})^*)^{-1}\right) \quad \text{ossia} \quad \mathrm{cl}\,\mathcal{G}(-A^{-1}) = [\mathcal{G}\left(((A^{-1})^*)^{-1}\right)]^\perp .$$

Usando la (3.58a) le uguaglianze precedenti si completano come segue:

$$[\mathcal{G}\left(((A^{-1})^*)^{-1}\right)]^\perp = \mathrm{cl}\mathcal{G}(-A^{-1}) = [\mathcal{G}(A^*)]^\perp .$$

Grazie al fatto che gli operatori aggiunti sono chiusi si trova

$$\mathcal{G}(A^*) = \mathcal{G}\left(((A^{-1})^*)^{-1}\right) \text{ ossia } A^* = \left((A^{-1})^*\right)^{-1}$$

da cui

$$(A^*)^{-1} = (A^{-1})^*$$

come volevamo provare. ∎

Ci interessa esaminare il caso in cui gli inversi sono continui. Si ha:

Teorema 284 *Valgono le due affermazioni seguenti:*

a) *l'operatore A verifichi le ipotesi del Teorema 283. Se A^{-1} è continuo allora* $\mathrm{dom}\, A^*$ *è denso in K e* $\mathrm{im}A^* = \mathrm{dom}(A^*)^{-1} = H$. *Inoltre, $(A^*)^{-1}$ è continuo.*

b) *sia A: $H \mapsto K$ un operatore lineare con dominio denso. Se A^* ha dominio denso in K ed immagine densa in H e se $(A^*)^{-1}$ è continuo allora anche A^{-1} è continuo ed ha dominio denso in K.*

Dim. Proviamo la proprietà **a)**. L'operatore A^{-1} è definito su $\mathrm{im}A$, denso in K ed essendo continuo esso ammette estensione continua a K. Sia C tale estensione. Dunque, $C \in \mathcal{L}(K, H)$ e quindi $C^* \in \mathcal{L}(H, K)$: *l'operatore C^* è continuo con dominio H.*

Si noti che C è la minima estensione chiusa di A^{-1} e quindi, dai teoremi 279 e 283,

$$C^* = (A^{-1})^* = (A^*)^{-1} \in \mathcal{L}(H, K)\,.$$

In particolare, $(A^*)^{-1}$ è continuo e $\mathrm{im}A^* = \mathrm{dom}(A^*)^{-1} = H$.

Proviamo la proprietà **b)**. L'ipotesi che A abbia dominio denso permette di definire A^*. Ciò fatto, per ipotesi, l'operatore A^*, che opera da K in H, ha le medesime proprietà elencate nel Teorema 283 per l'operatore A (che opera da H in K). Dunque:

1. $\mathrm{dom}\, A^{**}$ è denso in H perché A^{**} estende A;

2. $\mathrm{im}\, A^{**}$ è densa in K perché A^{**} estende A;

3. $(A^{**})^{-1} \in \mathcal{L}(H, K)$ perché $(A^*)^{-1}$ esiste ed è continuo.

Segue da qui che $\mathrm{im}A$ è densa in K. Infatti, da (3.58b),

$$(0, k_0) \in [\mathcal{G}(A)]^{\perp} = \mathcal{G}\left((-A^*)^{-1}\right) \implies k_0 = 0\,.$$

Usiamo ancora il fatto che A^{**} è la minima estensione chiusa di A e quindi, dalla proprietà 2 del Corollario 271, $(A^{**})^{-1}$ è la minima estensione chiusa di A^{-1}. Dunque, A^{-1} è continua perché $(A^{**})^{-1}$ è continua. ∎

3.6.3 Operatori Aggiunti e spettro

Sia ora A: $H \mapsto H$ un operatore lineare con dominio denso. I teoremi 282-284 in questo caso si riformulano come segue:

Lemma 285 *Sia A lineare da H in sé con dominio denso. Si ha $0 \in \sigma_p(A^*)$ se e solo se l'operatore A non ha immagine densa.*

Inoltre:

Teorema 286 *Sia A: $H \mapsto H$ lineare con dominio denso.*

$$0 \in \rho(A) \iff 0 \in \rho(A^*) \quad \text{ossia} \quad 0 \in \sigma(A) \iff 0 \in \sigma(A^*) \,.$$

Studiamo ora con maggiori dettagli le relazioni tra lo spettro di un operatore e quello del suo aggiunto. Dato che $\lambda \in \rho(A)$ equivale a $0 \in \rho(\lambda I - A)$ si vede che:

Teorema 287 *Vale: $\lambda \in \rho(A)$ se e solo se $\bar{\lambda} \in \rho(A^*)$; $\lambda \in \sigma(A)$ se e solo se $\bar{\lambda} \in \sigma(A^*)$.*

Invece, le singole componenti dello spettro non si conservano. Si ha invece:

Teorema 288 *Sia A un operatore lineare da H in H, con dominio denso. Se $\lambda \in \sigma_p(A)$ allora $\bar{\lambda} \in \sigma_p(A^*) \cup \sigma_r(A^*)$; se $\lambda \in \sigma_r(A)$ allora $\bar{\lambda} \in \sigma_p(A^*)$; se $\lambda \in \sigma_c(A)$ allora $\bar{\lambda} \in \sigma_c(A^*) \cup \sigma_r(A^*)$.*

Dim. La dimostrazione della seconda proprietà è immediata: se $\lambda \in \sigma_r(A)$ allora esiste $h \perp \text{im}\,(\lambda I - A)$ e per esso

$$0 = \langle h, (\lambda I - A)x \rangle = \langle (\bar{\lambda} I - A^*)h, x \rangle \qquad \forall x \in \text{dom}\, A \,.$$

E quindi $(\bar{\lambda} I - A^*)h = 0$.

Proviamo la prima. Se $\lambda \in \sigma_p(A)$ allora esiste $x_0 \neq 0$ per cui

$$0 = \langle (\lambda I - A)x_0, h \rangle = \langle x_0, (\bar{\lambda} I - A^*)h \rangle \qquad \forall h \in \text{dom}\, A^* \,.$$

Ciò vuol dire che $\text{im}\,(\bar{\lambda} I - A^*)$ non è densa e quindi, se $\bar{\lambda} \notin \sigma_p(A^*)$ deve essere $\bar{\lambda} \in \sigma_r(A^*)$.

Sia ora $0 \in \sigma_c(\lambda I - A)$. In questo caso $(\lambda I - A)^{-1}$ esiste su un dominio denso ma non è continuo e quindi $(\bar{\lambda} I - A^*)^{-1}$ esiste, definito su $\text{im}(-\bar{\lambda} I - A^*)$, e non è continuo, si veda il Teorema 287. Se $(\bar{\lambda} I - A^*)^{-1}$ non ha dominio denso in H allora $\bar{\lambda} \in \sigma_r(A^*)$, altrimenti $\bar{\lambda} \in \sigma_c(A^*)$. ∎

La situazione è riassunta nella tabella seguente.

$\lambda \in \rho(A)$	$\Longleftrightarrow$	$\bar{\lambda} \in \rho(A^*)$
$\lambda \in \sigma(A)$	$\Longleftrightarrow$	$\bar{\lambda} \in \sigma(A^*)$
$\lambda \in \sigma_p(A)$	$\Longrightarrow$	$\bar{\lambda} \in \sigma_p(A^*) \cup \sigma_r(A^*)$
$\lambda \in \sigma_r(A)$	$\Longrightarrow$	$\bar{\lambda} \in \sigma_p(A^*)$
$\lambda \in \sigma_c(A)$	$\Longrightarrow$	$\bar{\lambda} \in \sigma_c(A^*) \cup \sigma_r(A^*)$

Un corollario interessante del teorema (288) è:

Corollario 289 *Si sappia che $\sigma(A)$ è reale e che $A^* = A$. In tal caso $\sigma_r(A) = \emptyset$.*

Dim. Infatti, se $\lambda \in \sigma_r(A)$ ed è reale allora deve aversi anche $\lambda = \bar{\lambda} \in \sigma_p(A^*) = \sigma_p(A)$. Ciò è impossibile perché le tre componenti dello spettro sono disgiunte. ∎

E' importante sapere che il corollario precedente contiene un'ipotesi ridondante. Infatti

Teorema 290 *Se $A = A^*$ allora $\sigma(A)$ è reale.*

La dimostrazione è posposta nel paragrafo 3.6.6.

Gli operatori per cui $A = A^*$ si chiamano AUTOAGGIUNTI e sono importantissimi nelle applicazioni.

E' bene notare che la condizione $A = A^*$ in particolare richiede *l'uguaglianza dei domini.* Se invece A^* *estende* A, senza che si abbia l'uguaglianza dei domini, l'operatore A si chiama SIMMETRICO. Dunque, *ogni operatore autoaggiunto è simmetrico e inoltre $A \in \mathcal{H}$ è simmetrico se e solo se è autoaggiunto.*

Il risultato seguente mostra una proprietà importante degli operatori simmetrici:

Teorema 291 *Sia A: $H \mapsto H$ un operatore con dominio denso e simmetrico. Valgono le due proprietà seguenti:*

1. *Gli autovalori di A sono reali.*

2. *siano* λ *e* μ autovalori tra loro diversi. *Siano* x *ed* y *autovettori rispettivamente di* λ *e di* μ*:*

$$Ax = \lambda x\,, \qquad Ay = \mu x\,. \tag{3.59}$$

Allora, $x \perp y$.

<u>Dim</u>. Proviamo l'asserto 1. Sia λ un autovalore di A. Se $\lambda = 0$ niente va provato. Sia quindi $\lambda \neq 0$ e sia x un autovettore:

$$Ax = \lambda x\,.$$

Il prodotto interno con x dei due membri di quest'uguaglianza dà

$$\bar{\lambda}||x||^2 = \langle x, \lambda x\rangle = \langle x, Ax\rangle = \langle Ax, x\rangle = \langle \lambda x, x\rangle = \lambda ||x||^2\,.$$

Semplificando $||x||^2$ si trova $\lambda = \bar{\lambda}$.

Per provare l'asserto 2 si faccia il prodotti interno della prima delle uguaglianze in (3.59) con y e della seconda con x e si sommi. Usando il fatto che λ e μ sono reali si trova:

$$(\lambda - \mu)\langle x, y\rangle = 0\,.$$

Dato che $\lambda \neq \mu$, deve essere $x \perp y$. ∎

Si noti che questa dimostrazione è quella che si usa per provare l'asserto analogo relativo alle matrici simmetriche.

Si è osservato che $A \in \mathcal{H}$ è simmetrico se e solo se è autoaggiunto. L'esempio seguente mostra un operatore illimitato che è simmetrico ma non autoaggiunto.

Esempio 292 Mostriamo un operatore simmetrico che non è autoaggiunto. Ricordando il Teorema 290, si vede che basta mostrare un esempio di operatore simmetrico il cui spettro contiene il punto i, non reale. Consideriamo per questo lo spazio $X = L^2(0, +\infty)$. Si verifica facilmente che l'operatore lineare da X in se stesso definito da

$$f \longrightarrow \int_0^t e^{t-s} f(s)\,\mathrm{d}s$$

non è continuo. Consideriamo ora l'operatore A così definito:

$$\operatorname{dom} A = \left\{x \in C^1(0, +\infty)\,,\ x(0) = 0 \text{ e } x' \in L^2(0, +\infty)\right\}\,, \qquad Ax = ix'\,.$$

Integrando per parti si vede che l'operatore è simmetrico. Infatti, siano x ed y in dom A:

$$\langle Ax, y\rangle = \int_0^{+\infty} ix'(s)\overline{y(s)}\,\mathrm{d}s = \lim_{t\to+\infty} ix(t)\overline{y(t)} - \int_0^{+\infty} ix(s)\overline{y'(s)}\,\mathrm{d}s = \langle x, Ay\rangle$$

perchè una funzione che appartiene ad $L^2(0,+\infty)$ insieme alla sua derivata ha limite nullo per $t\to+\infty$. Calcolando

$$(iI - A)x = f$$

si trova l'equazione

$$x' = x + if$$

la cui soluzione è

$$x(t) = \int_0^t e^{t-s}[if(s)]\,\mathrm{d}s$$

e quindi non dipende da f in modo continuo. ∎

3.6.3.1 Il caso degli operatori chiusi

Sia A un operatore lineare da H in sé, con dominio denso. Ricordiamo che $\lambda\in\rho(A)$ quando $(\lambda I - A)^{-1}$ ha dominio denso in H ed è continuo. L'inverso che figura in questa definizione è l'inverso sinistro, perchè non si richiede che $(\lambda I - A)$ abbia immagine *uguale* ad H e ciò è alquanto insoddisfacente. Per esempio, se $(\lambda I - A)^{-1}$ è continuo definito su un s.spazio denso, esso può estendersi per continuità ad H ma la relazione tra A e tale estensione non è chiara. Mostriamo che questa difficoltà scompare se l'operatore A è chiuso:

Teorema 293 *Sia A chiuso e sia $\lambda\in\rho(A)$. In tal caso $(\lambda I - A)$ ha immagine* uguale *ad H.*

Dim. notando che $\lambda\in\rho(A)$ se e solo se $0\in\rho(\lambda I - A)$, basta provare che se $0\in\rho(A)$ allora im $A = H$.

L'operatore A^{-1} è chiuso perché A è chiuso, ed è continuo. Dunque il suo dominio è un s.spazio chiuso. Essendo $0\in\rho(A)$ segue che il dominio di A^{-1} è anche denso in H e quindi è H. ∎

Una proprietà analoga vale anche per i nuclei: Proviamo ora:

Teorema 294 *Sia A un operatore chiuso.* ker$(\lambda I - A)$ *è chiuso per ogni $\lambda\in\mathbb{C}$.*

In particolare, se A è lineare da H in K con dominio denso allora il s.spazio ker A^* *è chiuso in K.*

Dim. Basta provare l'asserto quando $\lambda = 0$. Sia A chiuso e sia (x_n) una successione in $\ker A$. Se essa non converge niente dobbiamo fare. Se però $x_n \to x_0$ allora dobbiamo provare che $x_0 \in \ker A$. Ciò è ovvio perché $((x_n, Ax_n)) = (x_n, 0) \to (x_0, 0)$ in $\mathcal{G}(A)$ ed il grafico è chiuso. Dunque, $x_0 \in \mathrm{dom} A$ ed $Ax_0 = 0$.

Questa osservazione si applica agli operatori aggiunti che sono sempre operatori chiusi. ∎

3.6.4 Spazi di Sobolev ed operatori aggiunti

Per semplicità lavoriamo con funzioni di una variabile definite su un intervallo (a, b), ma ciò che andiamo a dire si estende facilmente a funzioni di più variabili, a parte una precisazione importante, già notata al Paragrafo 2.3.1, e che ora ripeteremo.

Abbiamo detto che una funzione $f(s)$ appartiene a $W^{1,2}(a, b)$ quando è un elemento di $L^2(a, b)$ e inoltre esiste la sua *derivata in senso debole* e questa è una funzione di $L^2(a, b)$. Questa è la precisazione importante da fare: se la funzione dipende da più variabili ovviamente si considerano tutte le derivate parziali prime, in senso debole. Non è detto che le derivate parziali debbano esistere come limite del rapporto incrementale, nemmeno quasi ovunque; se invece si lavora con una funzione di *una sola* variabile allora se $f \in W^{1,2}(a, b)$ essa è derivabile q.o. in senso usuale e la sua derivata appartiene a $L^2(a, b)$. In tal caso la derivata calcolata nel senso delle distribuzioni coincide con la derivata usuale.[11]

Consideriamo ora l'operatore A da $L^2(a, b)$ in sé, il cui dominio sono le funzioni ϕ di classe C^∞ e a supporto compatto[12] su (a, b) e definito da:

$$(A\phi)(s) = \phi'(s)\,.$$

Calcoliamo $\mathrm{dom}\, A^*$. Una $x \in L^2(a, b)$ è in $\mathrm{dom}\, A^*$ se e solo se la trasformazione

$$x \longrightarrow \int_a^b x(s)\overline{\phi'(s)}\,\mathrm{d}s$$

[11] più correttamente, la derivata calcolata come limite del rapporto incrementale appartiene alla classe di equivalenza della derivata debole. Va notato che l'implicazione inversa non vale: esistono funzioni continue derivabili q.o. con derivata $L^2(a, b)$ che non appartengono a $W^{1,2}(a, b)$. Un esempio è la scala di Cantor.

[12] ossia nulle per $x \notin [c, d] \subseteq (a, b)$.

è continua su $L^2(a,b)$; ossia se e solo se essa si rappresenta come prodotto interno:

$$\int_a^b x(s)\overline{\phi'(s)}\,\mathrm{d}s = \int_a^b g(s)\overline{\phi(s)}\,\mathrm{d}s \qquad \forall \phi \in \operatorname{dom} A\,. \tag{3.60}$$

In tal caso, $x \in W^{1,2}(a,b)$ e la sua derivata, in senso debole, è $-g$.

Viceversa, sia $x \in W^{1,2}(a,b)$ e sia $-g$ la sua derivata in senso debole. Allora, la (3.60) vale per ogni ϕ di classe C^∞ ed a supporto compatto in (a,b). Dunque, $x \in \operatorname{dom} A^*$ se e solo se $x \in W^{1,2}(a,b)$. Si può quindi dire che lo spazio $W^{1,2}(a,b)$ si introduce per identificare il dominio degli operatori aggiunti di operatori di derivazione.

Analogamente, gli spazi di Sobolev $W^{k,2}(a,b)$ si introducono per descrivere il dominio dell'operatore A_k^*, con

$$(A_k\phi)(s) = \phi^{(k)}(s)$$

ancora definito sulle funzioni ϕ di classe C^∞ e a supporto compatto su (a,b).

3.6.5 Ortogonalità ed operatori aggiunti

In questo paragrafo illustriamo le proprietà intercorrenti tra nucleo ed immagina di un operatore lineare A e del suo aggiunto.

Due semplici osservazioni sono le seguenti:

Lemma 295 *Sia $A \in \mathcal{L}(H)$ e sia X_0 invariante per A: $AX_0 \subseteq X_0$. Sia $X = \operatorname{cl} X_0$. Il s.spazio chiuso X è invariante per A.*

Dim. Sia infatti $x \in H$,

$$x = \lim x_n\,, \qquad x_n \in X_0\,.$$

Vale:

$$Ax = \lim Ax_n \in X$$

perché $Ax_n \in X_0$ per ogni n e $X = \operatorname{cl} X_0$. ∎

Lemma 296 *Sia $A \in \mathcal{L}(H)$ e sia X un s.spazio invariante per A: sia cioè $AX \subseteq X$. Allora, $X^\perp$ è invariante per A^*.*

Dim. Bisogna provare che $A^*X^\perp \subseteq X^\perp$. Sia per questo $h \in X^\perp$ e sia $x \in X$. Si ha:

$$\langle x, A^*h\rangle = \langle Ax, h\rangle = 0$$

perché $Ax \in X$ ed è $h \perp X$. Ciò vale *per ogni* $x \in X$ e quindi $A^*h \in X^\perp$. ∎

Il risultato principale che vogliamo provare è il seguente:

Teorema 297 *Sia A lineare da H in K, con dominio denso in H. Si ha:*

a) $[\mathrm{im}\, A]^\perp = \ker A^*$.

b) $[\ker A^*]^\perp = \mathrm{cl}\ (\mathrm{im}\, A)$.

c) $\mathrm{cl}\ (\ker A) \subseteq [\mathrm{im}\, A^*]^\perp$.

  c') *se l'operatore A è chiuso vale l'uguaglianza:* $(\ker A) = [\mathrm{im}\, A^*]^\perp$.

d) $[\ker A]^\perp \supseteq \mathrm{cl}\ (\mathrm{im}\, A^*)$.

  d') *se l'operatore A è chiuso vale l'uguaglianza:* $[\ker A]^\perp = \mathrm{cl}\ (\mathrm{im}\, A^*)$.

Dim. La **a)** è il caso particolare in cui $H = K$ del Lemma 285 e la **b)** segue dalla proprietà **a)** passando agli ortogonali.

Proviamo che vale **c)**. Ricordiamo la (3.54b):

$$[\mathcal{G}(A^*)]^\perp = \mathrm{cl}\,\{(-Ah, h)\ h \in \mathrm{dom} A\}$$

e notiamo che basta provare $(\ker A) \subseteq [\mathrm{im}\, A^*]^\perp$ perché ogni ortogonale è un s.spazio chiuso.

Sia quindi $h_0 \in \ker A$. Dobbiamo provare $h_0 \perp \mathrm{im}\, A^*$. La condizione $h_0 \in \ker A$ si scrive anche $(0, h_0) \in \mathrm{cl}\,\{(-Ah, h)\ h \in \mathrm{dom} A\} = [\mathcal{G}(A^*)]^\perp$ dalla (3.54b). Dunque

$$\langle 0, k\rangle_K + \langle h_0, A^* k\rangle_H \qquad \forall k \in \mathrm{dom} A^* \quad \text{ossia} \quad h_0 \in [\mathrm{im} A^*]^\perp .$$

Quindi vale **c)**.

Mostriamo ora che se A è chiuso allora vale l'uguaglianza. Ricordiamo che essendo A chiuso con dominio denso allora A^* è esso stesso chiuso con dominio denso e $A^{**} = A$. Inoltre ricordiamo, dal Teorema 294, $\ker A$ è un s.spazio chiuso.

Dobbiamo provare l'inclusione inversa:

$$[\mathrm{im} A^*]^\perp \subseteq \ker A\,.$$

Sia per questo $h_0 \perp \mathrm{im} A^*$ ossia $\langle h_0, A^* k\rangle_H = 0$ per ogni $k \in \mathrm{dom} A^*$. Essendo $\mathrm{dom} A^*$ denso in K quest'uguaglianza mostra che $h_0 \in \ker A^{**} = \ker A$, come volevamo provare.

La **d)** e la **d')** si trovano passando agli ortogonali delle proprietà in **c)** e rispettivamente in **c')**. ∎

Infine, notiamo che la proprietà **a)** del teorema 297 può riformularsi come segue:

Teorema 298 *Sia A un operatore lineare da H in K, con dominio denso in H. L'immagine di A è densa in K se e solo se* $\ker A^* = 0$.

Rimane il problema di sapere se la *surgettività* di A può caratterizzarsi mediante A^*, questione risolta dal risultato seguente:

Teorema 299 *Sia A un operatore lineare* chiuso *di H in K, con dominio denso in H. L'operatore A è suriettivo se e solo se esiste $m > 0$ tale che*

$$||A^*y|| > m||y|| \qquad \forall y \in \operatorname{dom} A^* . \tag{3.61}$$

Dim. La (3.61) mostra che $\ker A^* = 0$ e quindi l'immagine di A è densa; la (3.61) mostra anche che $(A^*)^{-1}$ è continuo e quindi anche A^{-1} è continuo. Dunque, l'operatore A ha immagine densa ed inverso continuo, ossia $0 \in \rho(A)$. Essendo A chiuso segue che la sua immagine è chiusa, si ricordi il Teorema 293. ∎

3.6.6 Dimostrazioni posposte

Dimostrazione del TEOREMA 281.

La proprietà **a)** è ovvia. Proviamo **b)**. Osserviamo che $(A+B)^*$ esiste perché

$$\operatorname{dom}(A+B) = (\operatorname{dom}A) \cap (\operatorname{dom}B) \text{ è denso in } H .$$

Sia $k \in (\operatorname{dom}A^*) \cap (\operatorname{dom}B^*)$. Per ogni $x \in \operatorname{dom}(A+B)$ vale

$$\langle (A+B)x, k\rangle = \langle Ax, k\rangle + \langle Bx, k\rangle = \langle x, A^*k\rangle + \langle x, B^*k\rangle .$$

Dunque, $x \to \langle (A+B)x, k\rangle$ è continuo e dunque $k \in \operatorname{dom}(A+B)^*$. Notato ciò, le uguaglianze precedenti implicano che

$$(A+B)^*k = A^*k + B^*k .$$

Dunque, $(A+B)^*$ estende A^*+B^*. Proviamo **c)**. Per ipotesi, AB ha dominio denso in H e quindi $(AB)^*$ esiste.

Sia $w \in \operatorname{dom}A^* \subseteq W$.Allora, per ogni $h \in \operatorname{dom}AB$ si ha

$$\langle ABh, w\rangle = \langle Bh, A^*w\rangle .$$

Dunque il funzionale $h \mapsto \langle ABh, w\rangle$ è continuo se e solo se $A^*w \in \operatorname{dom}B^*$ ossia se e solo se $w \in \operatorname{dom}B^*A^*$ e in tal caso si ha

$$\langle ABh, w\rangle = \langle h, B^*A^*w\rangle .$$

Dunque, quando $w \in \text{dom}B^*A^*$ si ha $(AB)^*w = B^*A^*w$ ossia $(AB)^*$ estende B^*A^*.

Gli asserti che riguardano gli operatori continui sono ovvi. ∎

Dimostrazione del TEOREMA 290.

Sia $\lambda = \alpha + i\beta \in \sigma(A)$. Si deve provare che $\beta = 0$. Per assurdo sia $\beta \neq 0$. Il Teorema 291 mostra che $\lambda \notin \sigma_p(A)$ e quindi $\lambda \in \sigma_r(A) \cap \sigma_c(A)$.

Per ogni $x \in \text{dom}\, A$ vale

$$\langle(\lambda I - A)x, x\rangle = \lambda\langle x, x\rangle - \langle Ax, x\rangle \tag{3.62}$$

Ma ora, essendo $A = A^*$, $\langle Ax, x\rangle$ è reale. Infatti,

$$\langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle}\,.$$

Dunque si trova

$$\overline{\langle(\lambda I - A)x, x\rangle} = \bar{\lambda}\langle x, x\rangle - \langle Ax, x\rangle\,. \tag{3.63}$$

Sottraendo (3.62) e (3.63),

$$-2i\beta||x||^2 = \overline{\langle(\lambda I - A)x, x\rangle} - \langle(\lambda I - A)x, x\rangle = -2i\left(\mathcal{I}m\ \langle(\lambda I - A)x, x\rangle\right)\,.$$

Passando ai moduli si vede che

$$|\beta|\cdot||x||^2 = |\mathcal{I}m\ \langle(\lambda I - A)x, x\rangle| \leq |\langle(\lambda I - A)x, x\rangle| \leq ||(\lambda I - A)x||\cdot||x||\,. \tag{3.64}$$

L'ipotesi fatta per assurdo che $|\beta| > 0$ implica che l'inverso sinistro di $(\lambda I - A)$ è continuo, si veda la (2.66). Dunque, se $\beta \neq 0$ deve essere $i\beta \in \sigma_r(A)$. Esiste quindi $\xi \neq 0$ tale che

$$0 = \langle\xi, (\lambda I - A)x\rangle \qquad \forall x \in \text{dom}\, A\,.$$

Ciò in particolare implica che $\xi \in \text{dom}\, A^* = \text{dom}\, A$. Dunque, con $x = \xi$ si ha dalla (3.64):

$$|\beta|\,||\xi||^2 \leq |\langle\xi, (zlI - A)\xi\rangle| = 0\,.$$

Essendo $\xi \neq 0$ si trova la contraddizione $\beta = 0$. Dunque deve essere $\sigma(A) \subseteq \mathbb{R}$, come volevamo provare. ∎

3.7 Convergenza debole in spazi di Hilbert

Si è visto che in uno spazio di Banach di dimensione infinita i compatti sono "rari". Lo stesso ovviamente avviene anche nel caso particolare di spazi di Hilbert. Nel caso degli spazi di Hilbert la verifica è immediata. Infatti:

Teorema 300 *Sia H uno spazio di Hilbert di dimensione infinita. La sfera $S = \{x \mid ||x|| = 1\}$, che è chiusa,* non *è compatta.*

Dim. Sia infatti $\{e_n\}$ un sistema ortonormale numerabile. La distanza tra e_n ed e_m con $n \neq m$ vale $\sqrt{2}$. Infatti

$$||e_n - e_k||^2 = \langle e_n - e_k, e_n - e_k \rangle = ||e_n||^2 + ||e_k||^2 = 2\,.$$

Dunque la successione $\{e_n\}$, che appartiene ad S, non contiene s.successioni convergenti. ∎

Di conseguenza nemmeno la palla $\{x \mid ||x|| \leq 1\}$ è compatta e quindi *nessuna* palla è compatta. D'altra parte, si sa dai corsi precedenti che la proprietà di compattezza è cruciale nello studio dei problemi di minimo. Infatti:

Teorema 301 (di Weierstrass) *Una funzione f* continua *su un insieme* compatto *K ammette sia minimo che massimo.*

Accenniamo alla dimostrazione, che dovrebbe essere nota: per provare l'esistenza del minimo, si costruisce una SUCCESSIONE MINIMIZZANTE (k_n) in K, ossia una successione tale che

$$\lim f(k_n) = \inf\{f(k) \mid k \in K\}\,.$$

Si usa la compattezza di K per estrarre una sottosuccessione (k_{n_r}) convergente a $k_0 \in K$; e la continuità di f per concludere

$$f(k_0) = \lim f(k_{n_r}) = \inf\{f(k) \mid k \in K\}\,. \tag{3.65}$$

Ciò prova che k_0 è punto di minimo.

Osservazione 302 Alla medesima conclusione si giunge se la funzione f, invece di essere continua, verifica

$$\lim f(k_{n_r}) \geq f(k_0)\,.$$

Ovviamente la disuguaglianza stretta non può valere, e quindi segue la (3.65); ossia segue che k_0 è punto di minimo. ∎

La considerazione nell'osservazione precedente suggerisce di definire[13]

Definizione 303 *Sia $f(x)$ definita su un insieme D di uno spazio di Hilbert H e sia $x_0 \in D$. La funzione si dice* SEMICONTINUA INFERIORMENTE *in x_0 se per ogni successione (x_n) in D, convergente ad x_0, si ha*

$$\lim f(x_n) \geq f(x_0)\,.$$

Se $-f(x)$ è semicontinua inferiormente in x_0, la funzione si chiama SEMICONTINUA SUPERIORMENTE *in x_0.* ∎

Ricapitolando, la proprietà di "continuità" che serve per provare la parte del Teorema di Weierstrass relativa ai minimi è la semicontinuità inferiore; quella che serve per la parte relativa ai massimi è la semicontinuità superiore.

E' facile vedere che:

- se un insieme è compatto in una topologia, tale rimane anche in topologie *meno fini;*
- se una funzione a valori reali è continua oppure semicontinua inferiormente rispetto ad una topologia, tale rimane anche in topologie *più fini.*

Dunque conviene introdurre topologie *meno fini* di quella della norma, in modo da avere più insiemi compatti; ma sufficientemente fini in modo tale che almeno opportune classi di funzioni continue rimangano, se non continue, almeno semicontinue inferiormente.

Noi ci limiteremo ad introdurre *concetti di convergenza* per una topologia meno fine di quella della norma. Non descriveremo invece la topologia.

Definizione 304 Sia (x_n) una successione in uno spazio di Hilbert H. Diciamo che (x_n) CONVERGE DEBOLMENTE ad x_0 se

$$\lim\langle y, x_n\rangle = \langle y, x_0\rangle \qquad \forall y \in H\,. \;∎$$

Per indicare la convergenza debole si usa uno dei due simboli

$$\text{w}-\lim x_n = x_0 \qquad \text{oppure} \qquad x_n \rightharpoonup x_0\,.$$

E' ovvio che *la convergenza in norma implica la convergenza debole.* Il viceversa, in dimensione infinita, *non vale.*

[13] si veda l'analoga definizione 138, data nel contesto degli spazi di Banach.

Esempio 305 Sia $H = L^2(0,2\pi)$. Si è visto che ogni $f \in L^2(0,2\pi)$ si rappresenta

$$f(x) = \sum_{k=-\infty}^{k=+\infty} f_k e^{ikx}\,, \qquad f_k = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) e^{-ikx}\,\mathrm{d}x = \langle f, e_k\rangle$$

con $e_k = e^{ikx}$. Il Teorema di Riesz-Fisher implica che

$$\lim f_k = 0\,.$$

Questo vuol dire che,

$$\mathrm{w}-\lim e_k = 0\,.$$

Ma,

$$||e_k||_2 = \sqrt{2\pi}$$

e quindi la successione (e_k) *non* tende a zero in norma. ∎

Si ha:

Teorema 306 *Per la convergenza debole vale il teorema di unicità del limite.*

Dim. Se

$$\mathrm{w}-\lim x_n = x_0 \qquad \text{ed anche} \qquad \mathrm{w}-\lim x_n = y_0$$

allora si ha

$$\langle y, (x_0 - y_0)\rangle = \lim\{\langle y, x_n\rangle - \langle y, x_n\rangle\} = 0 \quad \forall y \in H\,.$$

Scegliendo in particolare $y = x_0 - y_0$ si vede che $y_0 = x_0$. ∎

Esistono alcune relazioni importanti tra la convergenza debole e le proprietà che valgono in norma. Tra queste:

Teorema 307 *Ogni successione debolmente convergente è limitata in norma.*

La dimostrazione è al paragrafo 2.10, dove è anche provato che ogni funzione continua in norma e convessa è debolmente semicontinua inferiormente. Qui limitiamoci a provare un caso particolare di quest'ultima affermazione.

Teorema 308 *vale:*

$$\text{Se} \quad \mathrm{w}-\lim x_n = x_0 \quad \text{allora} \quad \liminf ||x_n||_H \geq ||x_0||_H\,.$$

Dim. Per ogni y, con $||y||_H \leq 1$, vale

$$\langle y, x_n \rangle \to \langle y, x_0 \rangle .$$

In particolare questa vale scegliendo $y = x_0/||x_0||$ e quindi per ogni n si ha

$$||x_n|| \geq \langle \frac{x_0}{||x_0||}, x_n \rangle \longrightarrow \langle \frac{x_0}{||x_0||}, x_0 \rangle = ||x_0|| . \blacksquare$$

Il Teorema 308 può riformularsi dicendo che *la norma di H è debolmente semicontinua inferiormente.*

Ricordiamo ora che una funzione $f(x)$ su H *a valori reali* si dice CONVESSA quando il suo dominio è un insieme convesso e quando per ogni coppia di punti x ed y del dominio e per ogni $\lambda \in [0,1]$si ha:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) .$$

Vale il seguente risultato (che riformula per il caso degli spazi di Hilbert il Teorema 162):

Teorema 309 *Sia $f(x)$ una funzione su H* a valori reali *che è* continua *e* convessa. *La funzione $f(x)$ è debolmente semicontinua inferiormente.*

Si verifica facilmente che la norma è una funzione convessa e quindi il Teorema 308 è un caso particolare del precedente.

Inoltre:

Teorema 310 *Sia $\{x_n\}$ una successione tale che*

- $x_n \rightharpoonup x_0$;
- $||x_n|| \longrightarrow ||x_0||$.

Allora, la successione $\{x_n\}$ converge ad x_0 in norma.

Dim. Dobbiamo provare che

$$||x_n - x_0|| \to 0 .$$

E'

$$0 \leq ||x_n - x_0||^2 = \langle x_n - x_0, x_n - x_0 \rangle = ||x_n||^2 - 2\Re e \langle x_n, x_0 \rangle + ||x_0||^2 .$$

Per ipotesi, $||x_n||^2 \to ||x_0||^2$ e $\langle x_n, x_0 \rangle \to ||x_0||^2$. Dunque il membro destro ha limite nullo e ciò implica l'asserto. $\blacksquare$

Un insieme A si dice SEQUENZIALMENTE DEBOLMENTE CHIUSO quando contiene i limiti deboli di tutte le successioni convergenti $\{a_n\}$ con $a_n \in A$. Proviamo ora:

Teorema 311 *Sia A un insieme sequenzialmente debolmente chiuso in H. Allora, A è anche chiuso in norma.*

Dim. Sia A chiuso rispetto alla convergenza debole. Sia (x_n) una successione in A, che converge *in norma* ad un x_0.

La convergenza in norma implica la convergenza debole, e quindi (x_n) *converge debolmente* ad x_0. Dato che A è debolmente chiuso, si ha $x_0 \in A$; e quindi A è chiuso in norma. ∎

L'esempio 305 mostra che la superficie di una sfera di $L^2(0, 2\pi)$, pur essendo chiusa in norma, non è debolmente chiusa. Combinando questo col teorema precedente si vede che che la topologia della convergenza debole è effettivamente *meno fine* di quella della norma.

Vale però:

Teorema 312 *Sia A un sottoinsieme* convesso *e* chiuso *rispetto alla norma. L'insieme A è anche* chiuso *rispetto alla convergenza debole.*

La dimostrazione si trova nel paragrafo 2.10.

Inoltre:

Teorema 313 (di Alaoglu **in spazi di Hilbert)** *Ogni successione* limitata *in uno* SPAZIO DI HILBERT *ammette s.successioni debolmente convergenti.*

La dimostrazione si trova nel paragrafo 2.10.

Ricordiamo che il Teorema di Alaloglu *NON* vale in un generico spazio di Banach, ma vale in spazi di Banach *che sono spazi duali,* rispetto ad un *diverso* tipo di convergenza, che si chiama "convergenza debole stella".

La dimostrazione si trova nel paragrafo 2.10.

Infine, il teorema di Riesz mostra che ogni spazio di Hilbert è isomorfo al suo duale, e quindi anche al suo BIDUALE ossia al duale del duale. Ciò prova il Teorema 165.

3.8 Operatori compatti

Completiamo il nostro studio degli operatori tra spazi di Hilbert studiandone una classe particolare, quella degli "operatori compatti". Quando utile per chiarezza, indicheremo con C un operatore compatto.

Siano H e K spazi di Hilbert e sia $C \in \mathcal{L}(H, K)$. Essendo C continuo, il suo nucleo è un s.spazio chiuso di H e inoltre la restrizione di C a $[\ker C]^{\perp}$ è

iniettiva. Se in particolare $[\ker C]^{\perp}$ ha dimensione finita allora anche $\operatorname{im} C$ è uno spazio di dimensione finita e lo studio di C si fa semplicemente lavorando tra spazi di dimensione finita. In particolare, esistono basi $\{e_1, \ldots, e_n\}$ di $[\ker C]^{\perp}$ ed $\{\epsilon_1, \ldots, \epsilon_n\}$ di $\operatorname{im} C$ tali che

$$Cx = \sum_{i=1}^{n} \langle x, e_i \rangle \epsilon_i$$

per ogni $x \in H$ (e non solo per ogni $x \in [\ker C]^{\perp}$).

Osservazione 314 Si ha quindi una "diagonalizzazione" di C, ma rispetto a basi *diverse.* Si noti che le basi possono essere diverse anche se $H = K$. Per esempio sia $H = K = \mathbb{R}^2$ e sia C rappresentato dalla matrice

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

rispetto alla base canonica. L'operatore C *non* è diagonalizzabile scegliendo una medesima base per rappresentare $\mathbb{R}^2$ sia come spazio di partenza che d'arrivo; se però si scelgono come e_1 ed e_2 gli elementi della base canonica e invece

$$\epsilon_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \epsilon_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

allora

$$C[x_1 e_1 + x_2 e_2] = x_1 \epsilon_1 + x_2 \epsilon_2 \,. \quad ∎$$

La classe degli operatori C il cui nucleo ha codimensione finita ha quindi proprietà ben particolari. Sfortunatamente essa è troppo piccola per le applicazioni. Una classe più vasta di operatori, che ha proprietà ancora ben particolari e che però si incontra in numerose applicazioni è quella degli OPERATORI COMPATTI. Per definizione, un operatore si dice *compatto* quando ogni insieme *limitato* di H è trasformato in un insieme *relativamente compatto nella topologia della norma* di K.

Naturalmente, per vedere se un operatore è compatto basta verificare che *una* sfera ha per immagine un insieme relativamente compatto.

Osservazione 315 Ricordiamo che ogni insieme relativamente compatto è limitato. Dunque la proprietà di trasformare limitati in relativamente compatti implica la limitatezza e quindi la continuità dell'operatore. ∎

Ricordiamo che una successione è compatta quando ogni sua s.successione ammette punti limite. Ovviamente:

Teorema 316 *Siano H e K due soazi di Hilbert. L'operatore $C \in \mathcal{L}(H, K)$ è compatto se e solo se trasforma* ogni *successione limitata di H in una successione* compatta *di K (con la topologia della norma).*

Chiaramente tutti gli operatori con nucleo di codimensione finita, ossia con immagine di dimensione finita, trasformano insiemi limitati in insiemi relativamente compatti e inoltre:

Teorema 317 *Sia (C_n) una successione di operatori compatti. Se*

$$C = \lim C_n$$

(il limite nel senso di $\mathcal{L}(K, H)$), allora C è compatto.

In particolare ciò vale se per ciascun C_n si ha:

$$\dim [\operatorname{im} C_n] = c_n < +\infty \, .$$

Dim. Proviamo il teorema nel caso generale in cui ogni operatore C_n è compatto.

Proviamo che *ogni* successione (x_n) limitata di H ha per immagine una successione (Cx_n) che ammette s.successioni convergenti nella topologia della norma di K. Usiamo il procedimento diagonale di Cantor: si consideri la successione

$$n \to C_1 x_n \, .$$

Questa ammette s.successioni convergenti, perché l'operatore C_1 è compatto. Indichiamo col simbolo $(x_{1,n})$ una s.successione di (x_n) per cui $(C_1 x_{1,n})$ converge. La s.successione $(x_{1,n})$ è limitata perché la successione (x_n) è limitata. Dunque $(C_2 x_{1,n})$ ammette una s.successione convergente che indichiamo col simbolo $(C_2 x_{2,n})$.

Proseguendo in questo modo si costruiscono successioni $(x_{r,n})$ tali che:

- $(x_{r,n})$ è s.successione di $(x_{r-1,n})$;
- per ogni fissato i, la successione (di indice n) $(C_i x_{i,n})$ è convergente.
- Dunque, $(C_j x_{i,n})$ è convergente per ogni indice $j < i$, perché $(x_{i,n})$ con $i > j$ è s.successione di $(x_{j,n})$.

Si consideri ora la tabella seguente.

$C_1x_{1,1}$	$C_1x_{1,2}$	$C_1x_{1,3}$	$C_1x_{1,4}$	$C_1x_{1,5}$	$\ldots$
$C_2x_{2,1}$	$C_2x_{2,2}$	$C_2x_{2,3}$	$C_2x_{2,4}$	$C_2x_{2,5}$	$\ldots$
$C_3x_{3,1}$	$C_3x_{3,2}$	$C_3x_{3,3}$	$C_3x_{3,4}$	$C_3x_{3,5}$	$\ldots$
$C_4x_{4,1}$	$C_4x_{4,2}$	$C_4x_{4,3}$	$C_4x_{4,4}$	$C_4x_{4,5}$	$\ldots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Notiamo che la successione diagonale $(x_{n,n})$ è s.successione di ciascuna $(x_{i,n})$ e quindi $(C_i x_{n,n})$ converge per ogni i.

Applichiamo l'operatore C alla successione $\{x_{r,r}\}$ e proviamo che la successione $(Cx_{r,r})$ è convergente. Scriviamo per questo

$$||Cx_{n,n} - Cx_{m,m}|| \leq ||Cx_{n,n} - C_r x_{n,n}|| + ||C_r x_{n,n} - C_r x_{m,m}||$$
$$+||C_r x_{m,m} - Cx_{m,m}|| \leq ||C - C_r||\{||x_{n,n}|| + ||x_{m,m}||\} + ||C_r x_{n,n} - C_r x_{mm}|| \, .$$

Per ipotesi, (x_n) è limitata,

$$||x_n|| < M \qquad \forall n \, .$$

Ancora per ipotesi, $C_r \to C$ in $\mathcal{L}(H,K)$ e quindi per ogni $\epsilon > 0$ esiste r_ϵ tale che

$$||C - C_r|| < \epsilon/4M \, .$$

Con questo valore di r fissato, si ha

$$||C - C_r||\{||x_{n,n}|| + ||x_{m,m}||\} < \epsilon/2 \, .$$

Il numero r è ormai fissato e si sa che $(C_r x_{n,n})$ converge. Dunque si trova N_ϵ tale che, per n, m maggiori di N_ϵ vale

$$||C_r x_{n,n} - C_r x_{mm}|| < \epsilon/2 \, .$$

Dunque la successione $(Cx_{n,n})$ è fondamentale e quindi convergente.

Ciò prova che la successione (Cx_n) è compatta in K, come volevamo. ∎

In particolare,

Corollario 318 *L'insieme degli operatori compatti è un s.spazio chiuso di* $\mathcal{L}(H,K)$.

Infatti, che l'insieme degli operatori compatti è un insieme chiuso discende dalla dimostrazione precedente. Che è un s.spazio si vede facilmente.

Si ricordi che un operatore lineare continuo trasforma limitati in limitati e compatti in compatti. Dunque, se C è compatto, la sua composizione, a destra o a sinistra, con un operatore continuo è un operatore compatto. Dunque:

Teorema 319 *Siano H, K, X ed Y spazi di Hilbert. Sia $C\colon H \mapsto K$ un operatore compatto e sia $A \in \mathcal{L}(K,W)$ e $B \in \mathcal{L}(Y,K)$. Gli operatori AC e CB sono compatti.*

Vale inoltre:

Teorema 320 *L'operatore $C \in \mathcal{L}(H,K)$ è compatto se e solo se C^* è compatto.*

Dim. Dato che $C = C^{**}$, basta provare che se C è compatto il suo aggiunto lo è.

Per assurdo, supponiamo che C^* *non* sia compatto. In tal caso esiste una successione (k_n) limitata in K, e tale che (C^*k_n) *non* ammette s.successioni convergenti in norma. Dunque, per ogni successione di indici (n_k) esiste almeno un $\epsilon > 0$ tale che

$$||C^*x_{n_k} - C^*x_{n_m}|| > \epsilon$$

per infiniti indici k ed m. Passando ad una ulteriore s.successione, non è restrittivo assumere che ciò avvenga per ogni k e per ogni m.

Sia ora $h_{k,m}$ con $||h_{k,m}|| = 1$ e tale che

$$\epsilon/2 \le \langle h_{k,m}, C^*x_{n_k} - C^*x_{n_m}\rangle = \langle Ch_{k,m}, x_{n_k} - x_{n_m}\rangle\,. \tag{3.66}$$

Per ipotesi, l'operatore C è compatto. Dunque, l'insieme $\{Ch_{k,m}\}$ o è finito o ammette punti di accumulazione. Nel primo caso esiste z_0 ed esiste una successione (k_r, m_r) per cui

$$Ch_{k_r,m_r} = z_0\,.$$

Nel secondo caso esiste una successione (k_r, m_r) per cui

$$\lim Ch_{k_r,m_r} = z_0\,.$$

Limitandoci a considerare tale successione, si ha, per r sufficentemente grande,

$$\langle z_0, x_{n_{k_r}} - x_{n_{m_r}}\rangle = \langle z_0 - Ch_{k_r,m_r}, x_{n_{k_r}} - x_{n_{m_r}}\rangle + \langle Ch_{k_r,m_r}, x_{n_{k_r}} - x_{n_{m_r}}\rangle > \epsilon/4$$

perché vale (3.66) e perché il primo addendo del membro destro tende a zero.

Ciò non può darsi perché la successione $(\langle z_0, x_{n_{k_r}} - x_{n_{m_r}}\rangle)$ è una successione limitata di *numeri,* e quindi deve avere s.successioni convergenti per il *teorema di Bolzano–Weierstrass.* ∎

3.8.1 Lo spettro degli operatori compatti

Consideriamo un operatore compatto C da uno spazio di Hilbert *di dimensione infinita* H in sé e studiamone lo spettro. Esponiamo i risultati, posponendo le dimostrazioni. Le dimostrazioni sono al paragrafo 3.8.4.

Essendo C continuo, il suo spettro è non vuoto e limitato. Si sa inoltre che

$$\lambda \in \sigma(C) \quad \Longrightarrow \quad |\lambda| \le ||C||\,.$$

Dunque, o lo spettro è finito oppure è dotato di punti di accumulazione. Mostriamo prima di tutto che $\sigma(C)$ *può essere finito:*

Esempio 321 Sia $H = L^2(0,1)$ e sia C l'operatore da H in sé definito da

$$(Ch)(t) = \int_0^t h(s)\,\mathrm{d}s\,.$$

E' noto che $\sigma(C) = \{0\}$, si veda l'Esempio 190. Mostriamo che C è compatto. Notiamo per questo che l'immagine di C contiene soltanto funzioni continue e che C è anche continuo da $L^2(0,1)$ in $C(0,1)$. Inoltre, ogni s.insieme compatto di $C(0,1)$ è anche un s.insieme compatto di $L^2(0,1)$. Dunque basta provare che è compatto l'operatore

$$C \;:\; L^2(0,1) \to C(0,1)\,, \qquad (Ch)(t) = \int_0^t h(s)\,\mathrm{d}s\,.$$

Come si è notato, è sufficiente provare che l'immagine della sfera unità di $L^2(0,1)$ è compatta in $C(0,1)$. La continuità di C mostra che l'immagine è limitata. La disuguaglianza

$$|(Ch)(r) - (Ch)(t)| \le \left|\int_r^t |h(s)|\,\mathrm{d}s\right| \le \sqrt{|t-r|}\left[\int_0^1 |h(s)|^2\,\mathrm{d}s\right]^{1/2}$$

mostra l'equicontinuità dell'immagine, e quindi la compattezzo per il *teorema di Ascoli–Arzelà.* ∎

Nell'esempio precedente, $0 \in \sigma(C)$. Ciò non per caso. Infatti vale

Teorema 322 *Sia H uno spazio di Hilbert* di dimensione infinita. *Se C è compatto, il suo spettro contiene il punto* 0.

Se lo spettro di C è infinito, esso è numerabile *ed ha* 0 *come* unico *punto di accumulazione.*

Il risultato seguente va sotto il nome di ALTERNATIVA DI FREDHOLM.

Teorema 323 *Se* $\lambda \neq 0$*, allora* $\mathrm{im}\,(\lambda I - C)$ *è chiusa. Inoltre,* $\lambda \in \rho(C)$ *oppure* $\lambda \in \sigma_p(C)$*.*

Ossia, gli elementi *non nulli* dello spettro sono *autovalori.* Invece, il punto 0, che appartiene sempre allo spettro se $\dim H = +\infty$, può essere o meno un autovalore: nel caso dell'operatore visto nell'esempio 321 si ha $0 \in \sigma_c(C)$.

Ad ogni autovalore si associano i corrispondenti autovettori, uno o più, e ad ogni autovettore si associa una CATENA DI JORDAN. E' questa una successione, oppure una sequenza finita, (x_n) di vettori tali che

$$Ax_0 = \lambda x_0\,, \qquad Ax_n = \lambda x_n + x_{n-1} \quad \text{se } n > 0\,.$$

Dunque, il primo elemento x_0 della catena è un autovettore relativo all'autovalore λ.

Lo spazio generato da tutti gli elementi di catene di Jordan che corrispondono all'autovettore λ si chiama AUTOSPAZIO GENERALIZZATO di λ.

Vale:

Teorema 324 *Gli autospazi generalizzati di autovalori non nulli hanno dimensione finita.*

Naturalmente, se 0 è l'unico punto dello spettro, o anche se lo spettro è finito, lo spettro darà poche informazioni sull'operatore. Il caso in cui lo spettro dà informazioni "complete" sull'operatore è il caso in cui le catene di Jordan costituiscono un sistema massimale in H o almeno in $[\ker C]^\perp$, perché in tal caso l'operatore può rappresentarsi mediante "blocchi di Jordan". Un caso in cui ciò avviene è quello degli *operatori compatti e autoaggiunti:*

Teorema 325 *Sia* C compatto e autoaggiunto *sullo spazio di Hilbert* H *di dimensione infinita. Esiste una famiglia ortonormale* $\{v_n\}$ *(finita o numerabile) di autovettori di* C*,*

$$Cv_n = \lambda_n v_n \qquad \lambda_n \neq 0\,,$$

tale che

$$Cx = \sum \lambda_n \langle x, v_n\rangle v_n$$

per ogni $x \in H$*.*

La famiglia $\{v_n\}$ *è massimale in* $[\ker C]^\perp$*.*

Questo risultato generalizza la diagonalizzazione delle matrici simmetriche: rispetto a una base di autovettori l'operatore C può "scriversi in forma diagonale".

Chiameremo questa la DIAGONALIZZAZIONE di C.

3.8.2 Operatori compatti tra spazi diversi. Valori singolari

Premettiamo quaste osservazioni: un operatore A *autoaggiunto* in uno spazio di Hilbert è POSITIVO se

$$\langle Ax, x\rangle \geq 0 \qquad \forall x \in \mathrm{dom}A\,.$$

Ovviamente, gli autovalori di un operatore positivo sono non negativi:

Lemma 326 *Gli autovalori di un operatore positivo sono non negativi.*

Dim. Sia infatti λ un autovalore dell'operatore positivo A e sia x un corrispondente autovettore.

La definizione di operatore positivo implica che tale operatore è anche autoaggiunto e quindi i suoi autovalori sono reali (si veda il Teorema 291).

Si ha:

$$0 \leq \langle Ax, x\rangle = \lambda\|x\|^2 \quad \text{e quindi } \lambda \geq 0\,. \ ∎$$

Da ora in poi stabiliamo quanto segue: *gli autovalori di un operatore positivo sono ordinati in modo decrescente. Se l'operatore è compatto (e quindi ciascun autovalore ha molteplicità finita) ciascun autovalore è ripetuto tante volte quanto è la sua molteplicità.*

Passiamo ora a studiare un operatore $C \in \mathcal{L}(H, K)$ (spazi di Hilbert) che è compatto. Si noti che non si esclude che sia $H = K$.

Si sa, dal Teorema 319, che i due operatori

$$CC^* \in \mathcal{L}(K)\,, \qquad C^*C \in \mathcal{L}(H)$$

sono compatti e si verifica immediatamente che sono autoaggiunti e quindi si rappresentano rispettivamente come

$$C^*Ch = \sum_{i=1}^{+\infty} m_i\langle h, v_i\rangle v_i\,, \qquad CC^*k = \sum_{i=1}^{+\infty} \mu_i\langle k, w_i\rangle w_i\,. \tag{3.67}$$

Naturalmente, m_i e μ_i sono gli autovalori *non nulli* rispettivamente di C^*C e di CC^* mentre v_i e w_i rappresentano corrispondenti autovettori normalizzati.

Si vede immediatamente che C^*C e di CC^* sono ambedue operatori positivi rispettivamente in H ed in K e quindi i loro autovalori sono non negativi. I coefficienti m_i e μ_i essendo gli autovalori *non nulli*, sono strettamente positivi.

E' inoltre immediato vedere che i numeri m_i (ricordiamo, tutti non nulli) coincidono con i numeri μ_i (ricordiamo: anch'essi non nulli). Infatti, sia $\mu \neq 0$ tale che

$$CC^*v = \mu v\,.$$

Essendo $\mu \neq 0$, C^*v non è 0 e applicando C^* ai due membri si trova

$$(C^*C)C^*v = \mu C^*v$$

e quindi il numero μ (non nullo) è uno degli m (non nulli). In modo analogo si vede che ciascuno degli m_i coincide con uno dei numeri μ_i.

Osservazione 327 Nelle rappresentazioni (3.67) figurano i soli autovalori non nulli, ed abbiamo provato che essi sono i medesimi per CC^* come per C^*C. E' però possibile che 0 sia nello spettro di uno solo di questi operatori, come accade se $H = \mathbb{R}^2$, $K = \mathbb{R}$ e $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$. ∎

Introduciamo i numeri *non nulli*

$$\sigma_i = \sqrt{m_i}\,.$$

che si possono anche ottenere a partire dai μ_i e che si chiamano i VALORI SINGOLARI di C.

Generalmente si assume di ordinare i valori singolari in modo *non crescente,* ossia in modo compatibile con l'ordinamento scelto per gli autovalori.

Indichiamo con ω_i il vettore

$$\omega_i = \frac{1}{\sigma_i} C v_i$$

(si ricordi che i valori singolari sono non nulli.)

Vale:

Lemma 328 *L'insieme $\{\omega_i\}$ è ortonormale in K.*

Dim. Infatti,

$$\langle \omega_r, \omega_s \rangle = \langle \frac{1}{\sigma_r} C v_r, \frac{1}{\sigma_s} C v_s \rangle = \frac{1}{\sigma_r \sigma_s} \langle C^* C v_r, v_s \rangle = \frac{1}{\sigma_r \sigma_s} m_r \langle v_r, v_s \rangle$$

nullo se $r \neq s$ perché $v_r \perp v_s$, altrimenti uguale a 1. ∎

Poiché i $\{v_i\}$ sono un sistema ortonormale massimale in $[\ker C^*C]^\perp$, si può scrivere

$$x = \sum_{i=1}^{+\infty} \langle x, v_i \rangle v_i + n\,, \qquad n \in \ker C = \ker C^*C$$

e quindi

$$Cx = \sum_{i=1}^{+\infty} \langle x, v_i \rangle C v_i = \sum_{i=1}^{+\infty} \sigma_i \langle x, v_i \rangle \omega_i\,. \tag{3.68}$$

In particolare ciò mostra una "diagonalizzazione" per operatori compatti tra spazi diversi (in particolare, operanti nello stesso spazio, ma rispetto a basi diverse).

La (3.68) mostra che Cx è somma della serie, per ogni x fissato. In realtà vale di più: indichiamo con C_N l'operatore definito da

$$C_N x = \sum_{i=1}^{N} \sigma_i \langle x, v_i \rangle \omega_i \,.$$

Si ha

Lemma 329 *La successione di operatori (C_N) converge a C in $\mathcal{L}(H,K)$.*

Dim. Infatti, usando il fatto che la successione (σ_i) è non crescente,

$$||Cx - C_N x|| = \left|\left| Cx - \sum_{i=1}^{N} \sigma_i \langle x, v_i \rangle \omega_i \right|\right| = \left|\left| \sum_{i=N+1}^{+\infty} \sigma_i \langle x, v_i \rangle \omega_i \right|\right| \leq \sigma_{N+1} ||x||^2$$

e si sa che $\lim \sigma_{N+1} = 0$ *perché l'operatore C è compatto.* ∎

Dunque, ogni operatore compatto si approssima nella norma di $\mathcal{L}(H,K)$ mediante operatori con immagine di dimensione finita. Combinando ciò col teorema 317 si trova:

Teorema 330 *Un operatore $C \in \mathcal{L}(H,K)$ è compatto se e solo se è limite, in $\mathcal{L}(H,K)$, di una successione di operatori con immagine di dimensione finita.*

Osservazione 331 Gli operatori compatti possono definirsi anche in spazi di Banach e il teorema 317 vale anche in spazi di Banach. *Però in spazi di Banach esistono operatori compatti che non possono approssimarsi con operatori la cui immagine ha dimensione finita.* ∎

Per concludere, mostriamo una particolare rappresentazione sotto cui si possono porre gli operatori compatti da H in K.

Sia prima di tutto $C \in \mathcal{L}(H)$ compatto autoaggiunto e *positivo.* In tal caso si definisce

$$C^{1/2} x = \sum_{i=1}^{+\infty} \sqrt{\lambda_i} \langle x, v_i \rangle v_i \,.$$

Sia ora C compatto da H in K. Si definisce l'OPERATORE MODULO di C ponendo

$$|C| x = (C^* C)^{1/2} x = \sum_{i=1}^{+\infty} \sigma_i \langle x, v_i \rangle v_i \,.$$

Si noti che il simbolo $|C|$ *indica un operatore, e non un numero.*

Dato ora un generico operatore compatto A,

$$Ax = \sum_{i=1}^{+\infty} \sigma_i \langle x, v_i \rangle \omega_i$$

introduciamo l'operatore *(continuo ma generalmente non compatto)* come segue: se

$$x = \sum_{i=1}^{+\infty} \langle x, v_i \rangle v_i + n \,, \qquad n \in \ker A$$

poniamo

$$U_A x = \sum_{i=1}^{+\infty} \langle x, v_i \rangle \omega_i \,.$$

Le proprietà importanti di U_A sono:

- se $x \in [\mathrm{span}\,\{\, v_i \,\}]^\perp$ allora $U_A x = 0$;
- se $x \in \mathrm{cl}\,\mathrm{span}\,\{v_i\,\}$ allora

$$||U_A x||^2 = \left|\left| \sum_{i=1}^{+\infty} \langle x, v_i \rangle \omega_i \right|\right|^2 = \sum_{i=1}^{+\infty} |\langle x, v_i \rangle|^2 = ||x||^2 \,.$$

E' ora facile verificare che l'operatore compatto A si rappresenta come

$$A = U_A |A| \,.$$

Questa rappresentazione si chiama la RAPPRESENTAZIONE POLARE dell'operatore A.

3.8.2.1 Proprietà geometriche degli autovalori e dei valori singolari

Ricordiamo che la successione dei valori singolari di un operatore compatto è ordinata in modo non crescente.

Sia C compatto da H in K. Si è visto che

$$Cx = \sum \sigma_i \langle x, v_i \rangle \omega_i \,.$$

Dunque,

$$||Cx||^2 = \left|\left| \sum_{i=1}^{+\infty} \sigma_i \langle x, v_i \rangle \omega_i \right|\right|^2 = \sum_{i=1}^{+\infty} \sigma_i^2 |\langle x, v_i \rangle|^2 \leq \sigma_1^2 ||x||^2$$

e l'uguaglianza vale se $x = v_1$. Dunque,

Teorema 332 *Il numero σ_1, massimo valor singolare di C, è uguale a $||C||$. In particolare, se C è compatto ed autoaggiunto, $||C||$ è anche uguale a $\lambda_1 = \max\{\lambda_i\}$.*

Abbiamo così caratterizzato in modo geometrico il primo valor singolare ed il primo autovalore. Vogliamo estendere questa caratterizzazione ai successivi.

Consideriamo un operatore compatto autoaggiunto C, limitandoci a considerare il caso in cui tutti i suoi autovalori sono *non negativi.* In questo caso, *ordiniamo quelli strettamente positivi in modo non crescente,* $\lambda_i \geq \lambda_{i+1}$. Gli autovalori si elencano più volte quando ad essi corrispondono più autovettori linearmente indipendenti. Indichiamo con v_i un autovettore di norma 1 di λ_i, in modo da avere un sistema ortonormale $\{v_i\}$ massimale in $[\ker C]^\perp$.

Indichiamo con $\mathcal{L}[n]$ la famiglia di tutti i s.spazi di H di dimensione n. Un generico elemento di $\mathcal{L}[n]$ è

$$L = \operatorname{span}\{x_1, \dots, x_n\}.$$

Con $\tilde{L}_n$ indichiamo il particolare s.spazio generato dai primi n autovettori:

$$\tilde{L}_n = \operatorname{span}\{v_1, \dots, v_n\}.$$

Ricordiamo che stiamo studiando gli autovalori λ_i con $i > 1$. Un primo risultato è il seguente:

Lemma 333 *Sia C compatto autoaggiunto, con autovalori non negativi. Vale:*

$$\lambda_{n+1} = \max\left\{\langle Cx, x\rangle \mid ||x|| = 1\,,\ x \in [\tilde{L}_n]^\perp\right\}.$$

<u>Dim</u>. Se $x \in [\tilde{L}_n]^\perp$, si ha

$$x = \sum_{i=n+1}^{\infty} x_i v_i + n\,, \qquad x_i = \langle x, v_i\rangle\,, \qquad n \in \ker C$$

e quindi, essendo $||x||^2 = \sum |x_i|^2 + ||n||^2 = 1$,

$$\langle Cx, x\rangle = \sum_{i=n+1}^{\infty} \lambda_i x_i^2 \leq \lambda_{n+1}\left(\sum_{i=n+1}^{\infty} x_i^2\right) \leq \lambda_{n+1}\,.$$

In conclusione, per ogni $x \in [\tilde{L}_n]^\perp$ e di norma 1 si ha $\langle Cx, x\rangle \leq \lambda_{n+1}$. D'altra parte, se $x = v_{n+1}$ si ha $\langle Cv_{n+1}, v_{n+1}\rangle = \lambda_{n+1}$. Ciò prova il teorema. ∎

Il risultato precedente richiede l'esplicita conoscenza degli autospazi. In pratica interessano risultati che non fanno uso esplicito degli autospazi. Tra questi:

Teorema 334 *Sia C come nel Lemma 333. Si ha:*

$$\lambda_{n+1} = \min\left\{ \max\left\{ \langle Cx, x\rangle \;\mid ||x|| = 1\,,\; x \in L^{\perp}\right\}\,,\; L \in \mathcal{L}[n] \right\}.$$

<u>**Dim.**</u> Il lemma 333 mostra che

$$\lambda_{n+1} = \max\left\{ \langle Cx, x\rangle\,,\; \mid ||x|| = 1\,,\; x \in [\tilde{L}_n]^{\perp}\right\}\,.$$

Per provare il teorema basta mostrare che per ogni altra scelta di $L \in \mathcal{L}[n]$ si ha

$$\lambda_{n+1} \leq \max\left\{ \langle Cx, x\rangle\,,\; \mid ||x|| = 1\,,\; x \in L^{\perp}\right\}\,. \tag{3.69}$$

Sia $L = \mathrm{span}\,\{h_1\,, \ldots\,, h_n\}$ e sia

$$\phi = \sum_{i=1}^{n+1} \phi_i v_i\,.$$

Mostriamo che è possibile scegliere i coefficienti ϕ_i, non tutti nulli, in modo da avere $\phi \in L^{\perp}$. Ciò avviene se si trova una soluzione non nulla del sistema lineare

$$\sum_{i=1}^{n+1} \phi_i \langle h_j, v_i\rangle = 0 \qquad j = 1\,, \ldots\, n\,.$$

Questo è un sistema di n equazioni in $(n+1)$ incognite e quindi ammette soluzione non nulla. Si può quindi effettivamente trovare $\phi \perp L$ e, dividendo per $||\phi|| \neq 0$, si può assumere $||\phi|| = 1$. Per questo particolare elemento ϕ vale

$$\langle C\phi, \phi\rangle = \sum_{i=1}^{n+1} \lambda_i \phi_i^2 \geq \lambda_{n+1}||\phi||^2 = \lambda_{n+1}\,.$$

Dunque vale (3.69), come volevamo. ∎

Proviamo ora una caratterizzazione importante dei valori singolari. In questo caso C è compatto tra spazi di Hilbert H e K, può essere tra loro diversi.

Ricordiamo che i valori singolari per definizione sono non nulli ed ordinati in modo decrescente.

Con $\mathcal{A}[n]$ indichiamo la famiglia degli operatori lineari da H in K, ciascuno dei quali ha immagine di dimensione n *al più.*

Teorema 335 *Vale:*

$$\sigma_{n+1} = \min\{||C - A|| \;\mid A \in \mathcal{A}[n]\}\,.$$

Dim. Consideriamo l'operatore $\tilde{A}$ definito da

$$\tilde{A}x = \sum_{i=1}^{n} \sigma_i \langle x, v_i \rangle \omega_i \,.$$

Per quest'operatore si ha

$$(C - \tilde{A})x = \sum_{i=n+1}^{\infty} \sigma_i \langle x, v_i \rangle \omega_i$$

e quindi $||C - \tilde{A}|| = \sigma_{n+1}$. Ovviamente, $\tilde{A} \in \mathcal{A}[n]$.

Sia ora A un generico operatore che appartiene ad $\mathcal{A}[n]$. La sua restrizione a span $\{v_1\,, \dots\,, v_{n+1}\}$ non è iniettiva, perché A ha immagine di dimensione n al più, minore di quella del dominio. Dunque esiste $x = \sum_{i=1}^{n+1} x_i v_i$ tale che $Ax = 0$ e inoltre $||x|| = 1$. Per quest'elemento x vale

$$||(C - A)x|| = ||Cx|| \geq \sigma_{n+1}$$

e ciò completa la dimostrazione. ∎

3.8.3 Operatori compatti ed equazioni integrali di Fredholm

Consideriamo una funzione $K(t, s)$ continua su $[a, b] \times [a, b]$ e l'operatore da $L^2(a, b)$ in sé definito da

$$x \to Kx = \int_a^b K(t, s) x(s) \, \mathrm{d}s \,.$$

Si sa già che quest'operatore è continuo. Mostriamo che esso è addirittura compatto, usando il teorema di Ascoli–Arzelà. Ciò generalizza l'osservazione usata nell'Esempio 321.

Sia $x \in B$,

$$B = \{x \mid ||x|| < 1\} \,.$$

Se possiamo provare che KB è un insieme relativamente compatto di $L^2(a, b)$ allora K è un operatore compatto.

Si sa anche che K trasforma $L^2(a, b)$ in $C(a, b)$ e, come si è già visto, basta provare che KB è un s.insieme compatto di $C(a, b)$. La uniforme limitatezza di KB discende dalla continuità di K. Proviamo quindi l'equicontinuità.

Notiamo:

$$|(Kx)(t) - (Kx)(t')| = \left| \int_a^b [K(t,s) - K(t',s)] x(s) \, \mathrm{d}s \right|$$
$$\leq \left[\int_a^b |K(t,s) - K(t',s)|^2 \right]^{1/2} \left[\int_a^b |x(s)|^2 \right]^{1/2} \leq \left[\int_a^b |K(t,s) - K(t',s)|^2 \right]^{1/2} .$$

Sia $\epsilon > 0$. L'uniforme continuità di K mostra che esiste $\delta > 0$ tale che

$$|t - t'| < \delta \quad \Longrightarrow \quad |K(t,s) - K(t',s)| < \epsilon^2$$

e quindi, per $|t - t'| < \delta$,

$$|(Kx)(t) - (Kx)(t')| < \epsilon \, .$$

Ciò prova l'equicontinuità e quindi la compattezza.

Consideriamo ora l'equazione integrale

$$x = \mu K x + \phi = \mu \int_a^b K(t,x) x(s) \, \mathrm{d}s + \phi(t) \, .$$

Nel caso in cui K sia autoaggiunto, ossia nel caso in cui

$$K(t,s) = \overline{K}(s,t) \, ,$$

l'operatore K si può diagonalizzare rispetto ad un sistema ortonormale, mentre in generale si potrà scrivere

$$Kx = \sum_{i=1}^{+\infty} \sigma_i \langle x, v_i \rangle \omega_i \, .$$

Se accade che questa somma è finita, l'equazione integrale di Fredholm ha nucleo degenere; altrimenti, l'equazione integrale diviene

$$x(t) = \sum_{i=1}^{+\infty} \sigma_i \left[\int_a^b x(s) \bar{v}_i(s) \, \mathrm{d}s \right] \omega_i(t) + \phi(t) \, ,$$

forma che generalizza quella che abbiamo introdotto, per le equazioni con nucleo degenere, al paragrafo 1.1.3.

3.8.4 Dimostrazioni posposte

Si provano ora i teoremi relativi agli operatori compatti. Per questo avremo bisogno di introdurre alcune proprietà che valgono anche per operatori non compatti. Come si è già detto, per chiarezza indicheremo con A un generico operatore lineare e con C uno che è anche compatto.

Conviene seguire un'ordine un po' diverso da quello del paragrafo 3.8.1 e spezzare le dimostrazioni in vari lemmi.

Avremo spesso bisogno di lavorare con successioni *limitate* o addirittura *convergenti* (v_n) di elementi dell'immagine di $\lambda I - C$,

$$v_n = (\lambda I - C)x_n \,.$$

La successione (x_n) in generale non sarà né convergente né limitata. Però:

Lemma 336 *Sia* (v_n) *una successione limitata che appartiene ad* $\mathrm{im}(\lambda I - C)$*, con* C *compatto e* $\lambda \neq 0$*:*

$$v_n = \lim(\lambda I - C)x_n \,, \qquad \lambda \neq 0 \,. \tag{3.70}$$

Esiste una successione (k_n) limitata *e tale che*

$$v_n = (\lambda I - C)k_n \,.$$

Dim. Se (x_n) stessa è limitata, niente è da provare. Consideriamo il caso in cui l'uguaglianza (3.70) vale, con (x_n) successione *illimitata.*

Si rappresenti

$$X = \ker(\lambda I - C) \oplus [\ker(\lambda I - C)]^\perp \,, \qquad x_n = h_n + k_n \,.$$

Ovviamente, $(\lambda I - C)x_n = (\lambda I - C)k_n$ e quindi si può sostituire x_n con k_n. Basta dunque provare che la successione (k_n) è limitata.

Sia per assurdo la successione (k_n) illimitata. In tal caso,

$$y_n = \lim \frac{v_n}{||k_n||} = 0$$

perché (v_n) è limitata.

Usiamo l'ipotesi che l'operatore C è *compatto,* e la limitatezza di $(k_n/||k_n||)$, per estrarre dalla successione $\left(C\frac{k_n}{||k_n||}\right)$ una s.successione convergente. Cambiando nome agli indici, si può supporre

$$\lim C\frac{k_n}{||k_n||} = w_0 \,. \tag{3.71}$$

Essendo

$$(\lambda I - C)\frac{k_n}{||k_n||} = \frac{v_n}{||k_n||} \longrightarrow 0\,, \tag{3.72}$$

si ha anche

$$\lambda = \lim \frac{k_n}{||k_n||} = \lim \lambda \frac{k_n}{||k_n||} = \lim C\frac{k_n}{||k_n||} = w_0\,. \tag{3.73}$$

Questo ha due conseguenze:

1. Poiché λ *è non nullo* e $k_n/||k_n||$ ha norma 1, segue che[14]

$$||w_0|| = |\lambda| \neq 0 \qquad \text{ossia} \quad w_0 \neq 0\,. \tag{3.74}$$

2. Essendo $k_n \in [\ker(\lambda I - C)]^{\perp}$, segue che

$$w_0 \in [\ker(\lambda I - C)]^{\perp}\,. \tag{3.75}$$

D'altra parte, usando (3.71), (3.72), e (3.73) si ha

$$\begin{aligned}(\lambda I - C)\frac{w_0}{\lambda} &= \lim(\lambda I - C)\frac{k_n}{||k_n||} \\ &= \lim \lambda \frac{k_n}{||k_n||} - \lim C\frac{k_n}{||k_n||} = 0\end{aligned}$$

per (3.72) e (3.73). Dunque,

$$w_0 \in \ker(\lambda I - C)\,.$$

Questa e la (3.75) implicano $w_0 = 0$, in contrasto con (3.74).

La contraddizione trovata mostra che $\{k_n\}$ è limitata, come volevamo. ∎

Osservazione 337 Si noti che l'ipotesi della compattezza di C si è esplicitamente usata. *L'asserto precedente non vale per generici operatori.* ∎

Usiamo questo lemma per provare:

Teorema 338 *Se C è compatto e $\lambda \neq 0$ allora $(\lambda I - C)$ ha immagine chiusa.*

---

[14] si noti l'uso della continuità della norma: $\|\lim k_n/\|k_n\|\,\| = \lim \|\,k_n/\|k_n\|\,\|$.

Dim. Sia (v_n) una successione in $\mathrm{im}\,(\lambda I - C)$, convergente a v_0. Dobbiamo provare $v_0 \in \mathrm{im}\,(\lambda I - C)$ ossia che, per un opportuno x_0, si ha

$$v_0 = (\lambda I - C)x_0 \,.$$

Assumiamo quindi che sia $v_n = (\lambda I - C)x_n$. Come si è visto al Lemma 336, possiamo assumere che la successione (x_n) sia *limitata.* In questo caso, passando ad una s.successione che rinominiamo (x_n), si può assumere che (x_n) sia debolmente convergente e quindi che $\lim Cx_n = y$ in norma, perché C è compatto. Si ha quindi che

$$x_n = \frac{1}{\lambda}\left[v_n + Cx_n\right] \,, \qquad \lim x_n = \frac{1}{\lambda}[v_0 + y] \,,$$

ossia (x_n) converge *in norma* (e non solo debolmente) ad $x_0 = [v_0 - y]/\lambda$. D'altra parte C, essendo compatto, è *continuo.* Dunque, da $v_n = (\lambda I - C)x_n$, si ha $v_0 = (\lambda I - C)x_0$. Ciò prova che $v_0 \in \mathrm{im}\,(\lambda I - C)$, come volevamo. ∎

Proviamo ora:

Teorema 339 *Se l'operatore* $C \in \mathcal{L}(H, K)$ *è compatto e se lo spazio di Hilbert* K ha dimensione infinita, *allora* $0 \in \sigma(C)$.

Dim. Supponiamo che 0 sia nel *risolvente* di C. In questo caso, C^{-1} è continuo e quindi

$$I = CC^{-1}$$

è un operatore compatto. Dunque, la sfera $\{x \mid ||x|| \leq 1\}$ è compatta. Ciò non può essere se lo spazio di Hilbert H ha dimensione infinita, si veda il Teorema 57. ∎

Proviamo ora un lemma che sarà reso più preciso in seguito[15]:

Lemma 340 *Sia* C *compatto. Se* $\lambda \in \sigma(C)$ *non è zero, allora* $\lambda \in \sigma_p(C) \cup \sigma_r(C)$.

Dim. Sia $\lambda \neq 0$, e sia $\lambda \notin \sigma_p(C)$. Si è visto che l'immagine di $(\lambda I - C)$, con $\lambda \neq 0$, è chiusa. Se questa è diversa da X allora $\lambda \in \sigma_r(C)$. Se l'immagine è X allora, per il teorema di Banach, Teorema 104, $(\lambda I - C)^{-1}$ è continuo e quindi $\lambda \in \rho(C)$. ∎

[15]la dimostrazione usa il Teorema di Banach, Teorema 104: se un operatore lineare è continuo, invertibile e suriettivo, il suo inverso è continuo.

Osservazione 341 In particolare si è provato che lo spettro continuo di un operatore compatto, se non è vuoto, contiene il solo elemento 0. ∎

Notiamo che non abbiamo ancora provato che gli elementi non nulli di $\sigma(C)$ sono autovalori. Possiamo però provare:

Lemma 342 *L'insieme degli autovalori dell'operatore compatto C, se non è finito, ha per unico punto di accumulazione il punto* 0.

Dim. Ricordiamo che autovettori corrispondenti ad autovalori diversi sono linearmente indipendenti. Questo risultato, noto dai corsi di algebra lineare, è provato per completezza nel Lemma 345.

Essendo C continuo, il suo spettro è un insieme limitato e quindi se C ha infiniti autovalori, si trova una successione (λ_k) di autovalori tra loro diversi, che converge a λ_0. Supponiamo per assurdo che sia $\lambda_0 \neq 0$. Indichiamo con x_k un autovettore di λ_k di norma 1 e sia $X_n = \operatorname{span}\{x_1, \dots x_n\}$.

Lo spazio lineare X_n è trasformato in sé dall'operatore C,

$$CX_n \subseteq X_n$$

ed inoltre

$$(\lambda_n I - C)\, X_n \subseteq X_{n-1}$$

perché $(\lambda_n I - C)x_n = 0$.

Grazie al Lemma 345 (che proveremo successivamente), la dimensione di X_n è esattamente n e quindi $X_{n-1} \subseteq X_n$, l'inclusione essendo stretta. Dunque in X_n può trovarsi un vettore e_n di norma 1, che dista 1 da X_{n-1}. Mostriamo che se $\lambda_0 \neq 0$, la successione $\left(C\frac{e_n}{\lambda_n}\right)$ non ammette s.successioni convergenti. Ciò contrasta con la compattezza di C e mostra che $\lambda_0 = 0$. Per ottenere ciò, basta provare che per ogni n, m vale

$$\left|\left|C\frac{e_n}{\lambda_n} - C\frac{e_m}{\lambda_m}\right|\right| > 1\,.$$

Per fissare le idee, sia $n > m$ e si scriva

$$C\frac{e_n}{\lambda_n} - C\frac{e_m}{\lambda_m} = e_n - \left\{C\frac{e_m}{\lambda_m} + \left(I - \frac{C}{\lambda_n}\right)e_n\right\}\,.$$

I due vettori

$$\left(I - \frac{C}{\lambda_n}\right)e_n\,, \qquad C\frac{e_m}{\lambda_m}$$

appartengono a X_{n-1} mentre $e_n \in X_n$. Dunque,

$$\left\|C\frac{e_n}{\lambda_n} - C\frac{e_m}{\lambda_m}\right\| \geq \mathrm{dist}(e_n, X_{n-1}) \geq 1\,.$$

Si trova quindi che la successione $(C(e_n/\lambda_n))$ ha immagine priva di s.successioni convergenti; e la successione (e_n/λ_n) è limitata se $\lambda_n \to \lambda_0 \neq 0$. Ciò contrasta con la compattezza di C e mostra che deve essere $\lambda_0 = 0$. ∎

Vorremo provare che

$$\sigma_r(C) - \{0\} = \emptyset\,.$$

Per ora però proviamo:

Corollario 343 *Sia C compatto. L'insieme $\sigma_r(C)$ è finito oppure ha 0 come unico punto di accumulazione.*

Dim. Infatti, si sa dal teorema 288 che se $\lambda \in \sigma_r(C)$ allora $\bar{\lambda} \in \sigma_p(C^*)$ e C^* è compatto, si veda il Teorema 320; e quindi l'unico punto che può essere di accumulazione per $\sigma_r(C)$ è il punto 0. ∎

Per completare le dimostrazioni dei risultati relativi allo spettro di operatori compatti, dobbiamo far intervenire le proiezioni spettrali introdotte al Teorema 197.

Per i risultati già provati, si può trovare una successione decrescente di numeri positivi (r_n), $r_n \to 0$, tali che

$$\{\lambda\,:\ \ |\lambda| = r_n\} \subseteq \rho(C)\,.$$

Infatti, 0 è l'unico punto di accumulazione sia di $\sigma_p(C)$ che di $\sigma_r(C)$; e si è già visto che $\sigma_c(C) \subseteq \{0\}$.

Con un abuso di linguaggio comune nella teoria delle funzioni olomorfe, indichiamo con Γ_n la "curva" costituita dalle due circonferenze di centro 0 e di raggio rispettivamente r_n, r_{n+1}. Sia Ω_n la corona circolare delimitata da Γ_n. Consideriamo l'operatore

$$P_n = \frac{1}{2\pi i}\int_{\Gamma_n} (zI - C)^{-1}\,\mathrm{d}z\,.$$

Ricordiamo, dal Teorema 197:

- I due s.spazio $\mathrm{im}\,P_n$ ed $\mathrm{im}\,(I - P_n)$ sono complementari;
- L'insieme $\sigma_n = \sigma(C) \cap \Omega_n$ è lo spettro della restrizione di C ad $\mathrm{im}\,P_n$ (che è un s.spazio invariante per C).

Proviamo ora:

Teorema 344 *Sia C compatto. La proiezione P_n ha immagine di dimensione finita e quindi ogni elemento non nullo di $\sigma(C)$ è un* autovettore *il cui autospazio generalizzato ha dimensione finita.*

Dim. Notiamo che la funzione $1/z$ è olomorfa in Ω_n e quindi

$$P_n = \frac{1}{2\pi i}\int_{\Gamma_n}(zI-C)^{-1}\,\mathrm{d}z = \frac{1}{2\pi i}\int_{\Gamma_n}\left\{(zI-C)^{-1}-\frac{1}{z}\right\}\mathrm{d}z$$
$$= \frac{1}{2\pi i}\int_{\Gamma_n}(zI-C)^{-1}\frac{C}{z}\,\mathrm{d}z\,.$$

L'ultimo integrale si approssima *nella topologia di $\mathcal{L}(X)$*, mediante le somme di Riemann

$$P_n = \frac{1}{2\pi i}\left\{\sum_r (z_rI-C)^{-1}\frac{z_r-z_{r-1}}{z_r}\right\}C$$

(i punti z_r sono quelli di una partizione del sostegno di Γ_n). Per il Teorema 319, ciascuno degli operatori P_n è compatto e quindi anche P lo è, si ricordi il teorema 317.

Dunque, la palla di im P_n è compatta e quindi, per il Teorema 197, im P_n ha dimensione finita. L'insieme $\sigma_n = \sigma(C)\cap\Omega_n$ è lo spettro di della restrizione di C ad imP_n (chiamiamola C_n) ed imP_n ha dimensione finita. Segue che gli elementi di σ_n sono autovalori di C_n con autospazi di dimensione finita. Dunque sono anche autovalori di C e con i medesimi autospazi, quindi ancora di dimensione finita. ∎

Infine, per completezza, proviamo:

Lemma 345 *Siano $\lambda_1,\ldots,\lambda_k$ autovalori* distinti *di un operatore lineare A e sia x_k un autovettore di λ_k. Gli autovettori x_k sono linearmente indipendenti.*

Dim. Ricordiamo che gli autovettori, per definizione, sono non nulli. Dunque, in particolare $x_1 \neq 0$ così che l'insieme $\{x_1\}$ è linearmente indipendente e, se gli autovettori non sono linearmente indipendenti, esiste un *primo* n_0 per cui

$$x_{n_0+1} = \sum_{i=1}^{n_0}\alpha_i x_i\,. \tag{3.76}$$

Applicando l'operatore A ai due membri dell'uguaglianza si trova

$$\lambda_{n_0+1}x_{n_0+1} = \sum_{i=1}^{n_0}\alpha_i\lambda_i x_i\,. \tag{3.77}$$

Moltiplicando i due membri della (3.76) per λ_{n_0+1} e sottraendo la (3.77), si trova

$$\sum_{i=1}^{n_0}[\lambda_{n+1}-\lambda_i]\alpha_i x_i = 0\,.$$

Ciò mostra che x_{n_0+1} *non* è il *primo* degli autovettori linearmente dipendente dai *precedenti.* Ciò contraddice la scelta di n_0 e prova l'asserto. ∎

Osservazione 346 Si noti che l'asserto precedente vale *per ogni* operatore lineare A, anche non compatto ed anche non continuo. ∎

3.8.4.1 Il caso degli operatori compatti autoaggiunti

In questa parte useremo i lemmi 296 e 295, che ricapitoliamo:

- Sia $A\in\mathcal{L}(K)$ e sia X un s.spazio invariante per A: sia cioè $AX\subseteq X$. Allora, $X^\perp$ è invariante per A^*.
- Sia $A\in\mathcal{L}(K)$ e sia X_0 invariante per A: $AX_0\subseteq X_0$. Sia $X=\operatorname{cl}X_0$. Il s.spazio chiuso X è invariante per A.

Gli operatori compatti autoaggiunti godono della proprietà seguente:

Teorema 347 *Sia $C\in\mathcal{L}(H)$ un operatore compatto e autoaggiunto. Almeno uno dei due numeri $||C||$ oppure $-||C||$ appartiene a $\sigma(C)$.*

Proveremo in seguito questo teorema. Per ora illustriamone le conseguenze.

Una prima conseguenza è che il raggio spettrale di un operatore compatto autoaggiunto è uguale a $||C||$. In generale invece il raggio spettrale di un generico operatore lineare A è *minore della sua norma:* $r(A)\le||A||$. La disuguaglianza può essere stretta, anche se l'operatore è compatto (non autoaggiunto) come prova l'esempio della trasformazione da $\mathbb{R}^2$ in sé rappresentata, rispetto alla base canonica, dalla matrice

$$\left[\begin{array}{cc}0&1\\0&0\end{array}\right].$$

In questo caso lo spettro è $\{0\}$ mentre la norma dell'operatore è 1.

Notiamo ora che $||C||=0$ se e solo se $C=0$ e in questo caso $\sigma_p(C)=0$, ossia $\sigma_p(C)\neq\emptyset$. Se $C\neq0$,

$$\{+||C||\,,\ -||C||\}\cap\sigma(C)=\{+||C||\,,\ -||C||\}\cap\sigma_p(C)\,.$$

Corollario 348 *Se $C \in \mathcal{L}(H)$ è compatto autoaggiunto allora $\sigma_p(C) \neq \emptyset$. Se inoltre $C \neq 0$ allora C ha un autovalore non nullo.*

Assumiamo quindi esplicitamente $C \neq 0$. Il corollario precedente mostra che esistono autovalori non nulli dell'operatore C. Indichiamo con X_0 lo spazio lineare generato dagli autovettori di C *relativi ad autovalori non nulli.* Questo è un s.spazio di H invariante per C:

$$CX_0 \subseteq X_0 \qquad \text{con} \qquad X_0 = \operatorname{span}\{v_k \mid Cv_k = \lambda_k v_k\,,\ \lambda_k \neq 0\}\,.$$

Sia $X = \operatorname{cl} X_0$ così che X stesso è invariante per C.

Lemma 349 *Siano X_0 ed X gli spazi appena definiti. E': $X = \operatorname{cl} X_0 = [\ker C]^{\perp}$.*

Dim. Sia per assurdo $X \neq [\ker C]^{\perp}$ e quindi $X^{\perp} \neq \ker C$. Essendo X invariante per C, allora $X^{\perp}$ è invariante per C^* ed essendo $C = C^*$, $X^{\perp}$ è anch'esso invariante per C. La restrizione di C ad $X^{\perp}$ è essa stessa un operatore autoaggiunto e quindi ammette un autovalore, per il Corollario 348 e se $X^{\perp} \neq \ker C$ allora la restrizione di C ad $X^{\perp}$ è non nulla. Esiste quindi un autovalore di C un cui autovettore appartiene ad $X^{\perp}$. Ciò contrasta con la definizione di X che, per costruzione, contiene *tutti* gli autovettori di C *relativi ad autovalori non nulli.* La contraddizione trovata prova il teorema. ∎

Per costruzione, l'insieme degli autovettori (normalizzati) di C che corrispondono ad autovalori non nulli genera X_0 ed è quindi un sistema ortonormale massimale in X. Mostriamo:

Teorema 350 *Sia C compatto autoaggiunto. Si possono scegliere gli autovettori di C, di autovalore non nullo, in modo da avere un sistema ortonormale massimale di $X = [\ker C]^{\perp}$.*

Dim. Basta provare che gli autovettori si possono scegliere due a due ortogonali. Si è già visto al Teorema 291 che l'ortogonalità è automatica per autovettori che corrispondono ad autovalori diversi. Sia ora $\lambda_0 \neq 0$ un autovettore di molteplicità maggiore di 1 e sia N_0 il relativo autospazio. Essendo C compatto, la dimensione di N_0 è finita e quindi N_0 ammette una base ortonormale di autovettori perché la restrizione di C ad N_0 è un operatore autoaggiunto su uno spazio di dimensione finita. ∎

In conclusione, ogni $h \in H$ può rappresentarsi come

$$h = h_0 + \sum_{i=1}^{+\infty} \langle h, e_i \rangle e_i$$

con

$$Ce_i = \lambda_i e_i\,, \qquad \lambda_i \neq 0\,, \qquad Ch_0 = 0\,.$$

Dunque, per ogni $h \in H$ si ha anche

$$Ch = \sum_{i=1}^{+\infty} \lambda_i \langle h, e_i \rangle e_i\,.$$

Questa è la forma diagonale cercata dell'operatore C.

Passiamo ora a provare il Teorema 347. La dimostrazione richiede diversi passi. Proviamo prima di tutto due lemmi che valgono per operatori autoaggiunti, anche non compatti.

Lemma 351 *Siano x ed y in H e sia $A \in \mathcal{L}(H)$ un operatore autoaggiunto. Si ha:*

$$4\Re e \langle Ax, y \rangle = \langle A(x+y), (x+y) \rangle - \langle A(x-y), (x-y) \rangle\,.$$

Dim. Usando $A = A^*$, calcoliamo:

$$\begin{aligned}\langle A(x+y), (x+y) \rangle &= \langle Ax, x \rangle + \langle Ay, y \rangle + \langle Ax, y \rangle + \langle Ay, x \rangle \\ &= \langle Ax, x \rangle + \langle Ay, y \rangle + \langle Ax, y \rangle + \langle y, Ax \rangle \\ &= \langle Ax, x \rangle + \langle Ay, y \rangle + 2\Re e \langle Ax, y \rangle\,. \end{aligned} \tag{3.78}$$

Analogamente si vede che

$$\langle A(x-y), (x-y) \rangle = \langle Ax, x \rangle + \langle Ay, y \rangle - 2\Re e \langle Ax, y \rangle\,. \tag{3.79}$$

L'asserto segue sottraendo la (3.79) da (3.78). ∎

Proviamo ora:

Lemma 352 *Sia A un operatore lineare continuo ed autoaggiunto. Vale:*

$$||A|| = \sup \{ |\langle h, Ah \rangle|\,,\ ||h|| = 1 \}\,.$$

Dim. Sia

$$\alpha = \sup \{ |\langle h, Ah \rangle|\,,\ ||h|| = 1 \}\,.$$

Per ogni operatore lineare A vale $\alpha \leq ||A||$. Si deve provare che se A è autoaggiunto, allora la disuguaglianza non può essere stretta; ossia, si deve provare che se A è autoaggiunto, allora

$$||A|| \leq \alpha\,.$$

Ricordiamo, come conseguenza del Teorema di Riesz e del Teorema 122 che

$$||A|| = \sup\left\{\langle Ah, k\rangle\,,\ ||h|| = 1\,,\ ||k|| = 1\right\}\,.$$

D'altra parte, per il Lemma 351 e per l'identità del parallelogramma, essendo $||h|| = 1$, $||k|| = 1$,

$$\begin{aligned} &4\Re e\,\langle Ah, k\rangle \le |\langle A(h+k),(h+k)\rangle| + |\langle A(h-k),(h-k)\rangle| \\ &\le \alpha\left\{||h+k||^2 + ||h-k||^2\right\} = 2\alpha\left\{||h||^2 + ||k||^2\right\} = 4\alpha \end{aligned}$$

ossia

$$\Re e\,\langle Ah, k\rangle \le \alpha\,.$$

In generale, $\langle Ah, k\rangle$ è un numero complesso,

$$\langle Ah, k\rangle = e^{i\theta}\,|\langle Ah, k\rangle|\,.$$

Sostituendo h con $z = e^{-i\theta}h$ si trova

$$|\langle Ah, k\rangle| \le \alpha\,.$$

Questo calcolo può venir ripetuto per ogni h, k di norma 1, come volevamo provare. ∎

Osservazione 353 Come si è detto, per ogni operatore lineare A vale

$$||A|| = \sup\langle Ah, k\rangle\,. \tag{3.80}$$

L'estremo superiore si calcola al variare di h e di k in modo indipendente nella palla di raggio 1. Se A è autoaggiunto vale di più:

$$||A|| = \sup_{||h||=1} |\langle Ah, h\rangle|\,.$$

Si noti la presenza del modulo in quest'ultima uguaglianza.

E' indifferente mettere o meno il modulo nella (3.80). ∎

Torniamo ora a considerare un operatore compatto autoaggiunto C.

Dimostrazione del TEOREMA 347.

Come si è notato, si può assumere $C \neq 0$.

Sappiamo già che lo spettro dell'operatore C è reale perché C è autoaggiunto; e quindi

$$\sigma(C) \subseteq [-||C||, ||C||\,]\,.$$

Bisogna provare che uno almeno degli estremi di quest'intervallo appartiene allo spettro.

Si è provato nel Lemma 352 che

$$||C|| = \sup\{|\langle h, Ch\rangle|\,,\ ||h|| = 1\}\ .$$

Esiste quindi una successione (h_n), con $||h_n|| = 1$, tale che

$$\lim\langle Ch_n, h_n\rangle = \alpha \qquad \text{dove } \alpha = ||C|| \text{ oppure } \alpha = -||C||.$$

Si noti che $\alpha \neq 0$ perché si suppone $C \neq 0$.

Proviamo prima di tutto che

$$\lim[Ch_n - \alpha h_n] = 0\,. \tag{3.81}$$

Calcoliamo per questo

$$||Ch_n - \alpha h_n||^2 = ||Ch_n||^2 - 2\alpha\langle Ch_n, h_n\rangle + \alpha^2\,.$$

In questo calcolo si sono utilizzate le ipotesi che C è autoaggiunto, che α è reale e che $||h_n|| = 1$.

Si sa che $\lim\langle Ch_n, h_n\rangle = \alpha$ così che

$$\lim\left[-2\alpha\langle Ch_n, h_n\rangle + \alpha^2\right] = -\alpha^2\,.$$

Consideriamo ora la successione $(||Ch_n||^2)$. Vale

$$||Ch_n||^2 \leq \alpha^2\,.$$

Non supponiamo che la successione $\{||Ch_n||^2\}$ sia convergente. Però la dimostrazione precedente implica

$$0 \leq \limsup ||Ch_n - \alpha h_n||^2 = \limsup\left\{||Ch_n||^2 - 2\alpha\langle Ch_n, k_n\rangle + \alpha^2\right\} \leq 0\,.$$

Ciò prova (3.81).

La dimostrazione del Teorema 347 si completa come segue: essendo C compatto ed (h_n) limitata, esiste una s.successione (Ch_{n_r}) di (Ch_n), convergente in norma,

$$\lim Ch_{n_r} = k\,. \tag{3.82}$$

Usiamo ora il fatto che $\alpha \neq 0$ e notiamo che

$$h_{n_r} = \frac{1}{\alpha}\left\{Ch_{n_r} - [Ch_{n_r} - \alpha h_{n_r}]\right\}\ .$$

Si è visto che il termine in parentesi quadra tende a zero, mentre (Ch_{n_r}) tende a k. Dunque,

$$\lim h_{n_r} = \frac{1}{\alpha}k \qquad \text{e quindi} \qquad \lim Ch_{n_r} = C\left[\frac{1}{\alpha}k\right].$$

Di conseguenza, da (3.82),

$$C\left[\frac{1}{\alpha}k\right] = k \qquad \text{ossia} \qquad Ck = \alpha k\,.$$

Ciò prova che α è un'autovalore di C. Ricordando che α è $||C||$ oppure $-||C||$, si vede che l'asserto è provato. ∎

Capitolo 4

Distribuzioni e Trasformata di Fourier

In questo capitolo introduciamo i fatti essenziali della teoria delle "distribuzioni". Lo spazio lineare delle distribuzioni è dotato di una topologia la quale non proviene da una norma e quindi quest'argomento è sostanzialmente più complesso di quelli visti in precedenza. Non volendo entrare in queste complicazioni, ci limiteremo a definire la relazione di convergenza invece della topologia.

Una classe particolare di distribuzioni è quello delle "distribuzioni temperate". E' questo l'ambiente in cui è naturale definire la trasformata di Fourier. Le distribuzioni in generale vengono definite nel paragrafo 4.1 mentre le distribuzioni temperate e la trasformata di Fourier sono nel paragrafo 4.2. I due paragrafi sono scritti in modo da poter essere letti indipendentemente, con l'eccezione del paragrafo 4.3.2, nel quale si mostra che ogni distribuzione temperata è anche una distribuzione nel senso della definizione data al paragrafo 4.1.

4.1 Le distribuzioni

Prima di introdurre le distribuzioni, conviene vedere un esempio che ne giustifica l'uso. Consideriamo l'equazione a derivate parziali

$$x_t = x_s\,, \qquad x = x(t,s)\,. \tag{4.1}$$

Quest'equazione può scriversi nella forma

$$(1,-1)\cdot\nabla x(t,s) = 0$$

e ricordiamo che $(1,-1)\cdot\nabla x(t,s)$ è la derivata di $x(t,s)$ nella direzione di $t+s=0$. Consideriamo quindi una qualsiasi retta

$$t = -s + k$$

e la restrizione di $x(t,s)$ ad essa; ossia consideriamo la funzione $x(-s+k,s)$. Da

$$\frac{\mathrm{d}}{\mathrm{d}s}x(-s+k,s) = -x_t(-s+k,s) + x_s(-s+k,s) = 0$$

si vede che la funzione $x(-s+k,s)$ è costante. Naturalmente, il valore della costante dipende dalla retta su cui stiamo lavorando, ossia dipende da k:

$$x(-s+k,s) = \phi(k)\,.$$

Essendo $k = t+s$, la funzione $x(t,s)$ ha forma

$$x(t,s) = \phi(t+s) \tag{4.2}$$

per una "qualsiasi" funzione ϕ. Ovviamente però per poter giustificare i calcoli precedenti, dovremo richiedere che la ϕ sia almeno derivabile. Si capisce però che la (4.2) si potrà sostituire in (4.1) anche se ϕ non è ovunque derivabile, trovando che l'uguaglianza (4.1) vale "nei punti nei quali i calcoli hanno senso". Però, non è immediato capire quali proprietà più deboli la ϕ debba avere per giustificare i calcoli. Più ancora, ogni funzione continua è limite di funzioni di classe C^1 e quindi si può intendere la (4.2) come "soluzione in senso debole" di (4.1), indipendentemente dalla regolarità della ϕ e quindi senza preoccuparci dell'effettiva possibilità di sostituire la (4.2) nella (4.1).

Queste considerazioni, in quest'esempio semplicissimo, sono già un po' difficili da formalizzare. Ovviamente non si potranno estendere a casi più complessi. Mostriamo però un altro modo per *verificare* che la funzione $x(t,s) = \phi(t+s)$ è soluzione di (4.1). Moltiplichiamo ambedue i membri per una "funzione test" $\psi(t,s)$ di classe $C^\infty(\mathbb{R}^2)$ e integriamo per parti. Per evitare la comparsa di integrali impropri e dei limiti delle funzioni per t, s tendenti ad infinito, supponiamo che $\psi(t,s)$ abbia SUPPORTO COMPATTO, ossia che sia zero per t ed s abbastanza grandi. Integrando per parti si trova

$$\left\{ \begin{array}{l} \int_{\mathbb{R}^2} x_t(t,s)\psi(t,s)\,\mathrm{d}t\,\mathrm{d}s = -\int_{\mathbb{R}^2} x(t,s)\psi_t(t,s)\,\mathrm{d}t\,\mathrm{d}s\,, \\ \int_{\mathbb{R}^2} x_s(t,s)\psi(t,s)\,\mathrm{d}t\,\mathrm{d}s = -\int_{\mathbb{R}^2} x(t,s)\psi_s(t,s)\,\mathrm{d}t\,\mathrm{d}s \end{array} \right.$$

e quindi l'uguaglianza

$$-\int_{\mathbb{R}^2} x(t,s)\psi_t(t,s)\,\mathrm{d}t\,\mathrm{d}s = -\int_{\mathbb{R}^2} x(t,s)\psi_s(t,s)\,\mathrm{d}t\,\mathrm{d}s \tag{4.3}$$

è soddisfatta da tutte le soluzioni $x(t,s)$ di (4.1) che sono di classe C^1. Si noti però che la (4.3) non richiede nessuna regolarità alla $x(t,s)$. Interpretiamo quindi la (4.3) come una "forma debole" di (4.1) e ciò suggerisce di dire che una funzione $g(t,s)$ che verifica

$$-\int_{\mathbb{R}^2} x(t,s)\psi_t(t,s)\,\mathrm{d}t\,\mathrm{d}s = -\int_{\mathbb{R}^2} g(t,s)\psi(t,s)\,\mathrm{d}t\,\mathrm{d}s$$

è "derivata in senso debole" (rispetto a t) di $x(t,s)$. Analogamente si definisce la derivata in senso debole rispetto ad s. Si confronti con l'analoga definizione già incontrata nella definizione degli spazi di Sobolev.

La teoria delle distribuzioni nasce tentando di estendere le considerazioni precedenti.

4.1.1 Lo spazio $\mathcal{D}(\mathbb{R}^n)$ delle funzioni test

Si ricordi l'esempio di Cauchy di funzione di classe C^∞ non sviluppabile in serie di Taylor:

$$\phi(x) = \begin{cases} e^{-1/x^2} & x > 0 \\ 0 & x \leq 0\,. \end{cases}$$

Sia ora

$$\phi(x) = \begin{cases} e^{-1/(1-x^2)} & -1 < x < 1 \\ 0 & |x| \geq 1\,. \end{cases}$$

Questa è una funzione di classe $C^\infty(\mathbb{R})$ nulla per $|x| \geq 1$. A partire da questa funzione si costruiscono facilmente funzioni di classe C^∞ su $\mathbb{R}^n$, nulle per $||x|| \geq R$. Per esempio la funzione

$$\phi(x_1, x_2, \dots, x_n) = \begin{cases} e^{-1/(1-||x||^2/R^2)} & ||x||^2 = x_1^2 + \cdots x_n^2 < R^2 \\ 0 & ||x|| \geq R^2\,. \end{cases}$$

Ciò suggerisce:

Definizione 354 Sia ϕ definita su $\mathbb{R}^n$. Si chiama SUPPORTO di ϕ l'insieme

$$\operatorname{supp}\phi = \operatorname{cl}\{x \mid \phi(x) \neq 0\}\,.$$

Una funzione definita su $\mathbb{R}^n$, di classe C^∞ ed a supporto compatto si chiama una FUNZIONE TEST su $\mathbb{R}^n$. ▮

L'insieme delle funzioni test è ovviamente uno spazio lineare che si indica col simbolo $\mathcal{D}(\mathbb{R}^n)$. Nello spazio $\mathcal{D}(\mathbb{R}^n)$ si definisce una topologia la cui descrizione è piuttosto complessa. Limitiamoci a definire la convergenza di successioni:

Definizione 355 Sia (ϕ_n) una successione in $\mathcal{D}(\mathbb{R}^n)$ e sia $\phi \in \mathcal{D}(\mathbb{R}^n)$. Diciamo che

$$\lim \phi_n = \phi$$

se:

- Esiste un compatto K che contiene i supporti sia delle ϕ_n che di ϕ;
- vale $\phi_n(x) \to \phi(x)$ uniformemente su $\mathbb{R}^n$ ed inoltre anche ciascuna derivata parziale di ogni ordine di ϕ_n converge alla corrispondente derivata di ϕ, uniformemente su $\mathbb{R}^n$. ▮

L'enunciato della seconda proprietà è alquanto involuto. Per semplificarlo, conviene introdurre una notazione per le derivate parziali. Introduciamo prima di tutto la nozione di MULTIINDICE. Un multiindice α è una sequenza ordinata di n interi non negativi. Il numero n è uguale alla dimensione di $\mathbb{R}^n$. Dunque,

$$\alpha = (k_1, k_2, \ldots, k_n)\,, \qquad k_i \geq 0\,.$$

La LUNGHEZZA di α è

$$|\alpha| = k_1 + k_2 + \cdots + k_n\,.$$

Al multiindice α si associa l'operatore di derivata parizale

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{k^1} \partial x_2^{k_2} \cdots \partial x_n^{k_n}}\,.$$

Come al solito, la derivata di ordine zero lascia invariata la funzione. Per esempio, con $n = 2$ ed $\alpha = (0, 4)$, si trova

$$\frac{\partial^{0+4}}{\partial x_1^0 \partial x_2^4} \phi(x_1, x_2) = \frac{\partial^4}{\partial x_2^4} \phi(x_1, x_2)\,.$$

Con questa notazione, la seconda proprietà richiesta per la convergenza è

$$D^\alpha \phi_n \to D^\alpha \phi$$

uniformemente su $\mathbb{R}^n$ per ogni multiindice α.

Lo spazio $\mathcal{D}(\mathbb{R}^n)$ si intenderà sempre dotato della RELAZIONE DI CONVERGENZA sopra descritta.

Osservazione 356 Sia α un multiindice. La relazione di convergenza in $\mathcal{D}(\mathbb{R}^n)$ è stata definita in modo che D^α sia una trasformazione continua da $\mathcal{D}(\mathbb{R}^n)$ in sé nel senso che se $\phi_n \to \phi$ in $\mathcal{D}(\mathbb{R}^n)$ allora anche $D^\alpha \phi_n \to D^\alpha \phi$ in $\mathcal{D}(\mathbb{R}^n)$. ∎

4.1.2 Le distribuzioni

Si chiama DISTRIBUZIONE un funzionale lineare e continuo su $\mathcal{D}(\mathbb{R}^n)$; ossia, una distribuzione è un funzionale lineare χ su $\mathcal{D}(\mathbb{R}^n)$ tale che

$$\text{se } \phi_n \to 0 \text{ in } \mathcal{D}(\mathbb{R}^n) \text{ allora } \langle\!\langle \chi, \phi_n \rangle\!\rangle \to 0 \qquad .$$

Prima di mostrare le proprietà delle distribuzioni, vediamo alcuni esempi.

Esempio 357 Sia $f(x)$ una funzione su $\mathbb{R}^n$ che è localmente integrabile, ossia tale che

$$\int_K |f(x)| \, \mathrm{d}x < +\infty$$

per ogni compatto K.

Se $\phi_n \to 0$ in $\mathcal{D}(\mathbb{R}^n)$ allora esiste un compatto K che contiene i supporti di tutte le ϕ_n e quindi

$$\int_{\mathbb{R}^n} f(x)\phi_n(x) \, \mathrm{d}x = \int_K f(x)\phi_n(x) \, \mathrm{d}x \, .$$

Inoltre, $\phi_n \to 0$ uniformemente su K. Dunque si ha anche

$$\lim \int_{\mathbb{R}^n} f(x)\phi_n(x) \, \mathrm{d}x = \lim \int_K f(x)\phi_n(x) \, \mathrm{d}x = 0 \, ;$$

ossia, la trasformazione

$$\phi \to \langle\!\langle D_f, \phi \rangle\!\rangle = \int_{\mathbb{R}^n} f(x)\phi(x) \, \mathrm{d}x$$

è una distribuzione, che si chiama la DISTRIBUZIONE REGOLARE identificata da f. Per evitare ambiguità, noi indicheremo sempre col simbolo D_f la distribuzione regolare identificata da f ma è ovvio che in pratica questa verrà indicata semplicemente col simbolo f. ∎

Si noti che la condizione che i supporti delle ϕ_n siano contenuti in uno stesso compatto K quando $\phi_n \to 0$ è essenziale nell'esempio precedente.

Esempio 358 Si indica con δ la distribuzione che a $\phi \in \mathcal{D}(\mathbb{R}^n)$ associa $\phi(0)$,

$$\langle\!\langle \delta, \phi \rangle\!\rangle = \phi(0) \, .$$

La continuità del funzionale $\phi \to \phi(0)$ si verifica facilmente. Questa distribuzione si chiama la δ DI DIRAC (su $\mathbb{R}^n$). ∎

Esempio 359 Si fissi $\xi \in \mathbb{R}^n$ ed una direzione v in $\mathbb{R}^n$ e si consideri il funzionale che a $\phi \in \mathcal{D}(\mathbb{R}^n)$ associa

$$\lim_{\epsilon \to 0} \frac{\phi(\xi + \epsilon v/2) - \phi(\xi - \epsilon v/2)}{\epsilon} = [\nabla\phi(\xi)] \cdot v$$

La continuità su $\mathcal{D}(\mathbb{R}^n)$ del funzionale così costruito si verifica facilmente (si ricordi l'osservazione 356) e quindi questo funzionale è una distribuzione, che si chiama il DIPOLO in ξ di asse v. ∎

Esempio 360 Sia $f(x)$ una funzione definita su una superficie Σ, localmente integrabile su Σ. E' una distribuzione la trasformazione

$$\phi \to \int_\Sigma f(x)\phi(x)\,\mathrm{d}\Sigma\,.$$

Questa distribuzione, pur essendo definita da un integrale, *non è una distribuzione regolare* perchè l'integrale non è esteso a $\mathbb{R}^n$. In questo contesto, la funzione $f(x)$ si chiama *densità superficiale* (di massa, di carica, ...).

Questa distribuzione si chiama anche DISTRIBUZIONE DI SEMPLICE STRATO su Σ. ∎

Esempio 361 Sia $f(x)$ una funzione definita su una superficie Σ, localmente integrabile su Σ. E' una distribuzione la trasformazione

$$\phi \to \int_\Sigma f(x)\frac{\mathrm{d}}{\mathrm{d}n}\phi(x)\,\mathrm{d}\Sigma$$

dove $\mathrm{d}/\,\mathrm{d}n$ indica la derivata normale a Σ. Si interpretra questo numero come ottenuto da una distribuzione di dipoli su Σ, ciascuno di asse perpendicolare a Σ. Questa distribuzione si chiama DISTRIBUZIONE DI DOPPIO STRATO su Σ, di densità $f(x)$. ∎

4.1.3 Lo spazio $(\mathcal{D}(\mathbb{R}^n))'$

Col simbolo $(\mathcal{D}(\mathbb{R}^n))'$ si intende lo spazio lineare delle distribuzioni su $\mathbb{R}^n$, dotato della RELAZIONE DI CONVERGENZA seguente:

$$\chi_n \to \chi \qquad \text{quando } \textit{per ogni } \phi \text{ si ha} \qquad \langle\langle \chi_n, \phi\rangle\rangle \to \langle\langle \chi, \phi\rangle\rangle$$

(si confronti con la definizione di convergenza debole e debole stella).

Si provi per esercizio che le due trasformazioni seguenti sono continue da $(\mathcal{D}(\mathbb{R}^n))'$ *in sé:*

$$\chi \to \chi_0 + \chi\,, \qquad \chi \to \alpha\chi\,.$$

Esempio 362 Sia (h_n) un'identità approssimata. Per ogni $\phi \in \mathcal{D}(\mathbb{R})$ si ha

$$\lim_n \int_{-\infty}^{+\infty} h_n(t-s)\phi(s)\,\mathrm{d}s = \phi(t)\,,$$

si veda il paragrafo 2.4. Ciò vale in particolare per $t = 0$ e quindi

$$\lim_n \int_{-\infty}^{+\infty} h_n(-s)\phi(s)\,\mathrm{d}s = \phi(0)\,, \qquad \phi \in \mathcal{D}(\mathbb{R})\,.$$

Dunque, la successione di distribuzioni regolari definite da

$$\phi \to \int_{-\infty}^{+\infty} h_n(-s)\phi(s)\,\mathrm{d}s$$

converge in $(\mathcal{D}(\mathbb{R}))'$ alla δ di Dirac. Si dice più brevemente che "le identità approssimate approssimano la δ di Dirac". ∎

Nello spazio $(\mathcal{D}(\mathbb{R}^n))'$ possono definirsi altre operazioni. Tra queste:

- la TRASLAZIONE. Sia $h \in \mathbb{R}^n$ fissato. La traslazione $\mathcal{T}_h$ su $\mathcal{D}(\mathbb{R}^n)$ è definita da
$$[\mathcal{T}_h\phi](x) = \phi(x-h)\,.$$
Si verifica facilmente che questa trasformazione è lineare e continua da $\mathcal{D}(\mathbb{R}^n)$ in sé. Si estende questa trasformazione alle distribuzioni ponendo
$$\langle\langle \mathcal{T}_h\chi, \phi\rangle\rangle = \langle\langle \chi, \mathcal{T}_{-h}\phi\rangle\rangle\,.$$
Si noti che se $\chi = D_f$, una distribuzione regolare, questa definizione corrisponde alla formula di cambiamento di coordinate:
$$\int_{\mathbb{R}^n} f(x-h)\phi(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} f(x)\phi(x+h)\,\mathrm{d}x.$$

- l'OMOTETIA. Sia a un numero reale. L'omotetia $\mathcal{R}_a$ su $\mathcal{D}(\mathbb{R}^n)$ è definita da
$$[\mathcal{R}_a\phi](x) = \phi(ax)$$
e corrisponde a cambiare la scala del fattore a su ciascun asse coordinato. Si verifica facilmente che questa trasformazione è lineare e continua da $\mathcal{D}(\mathbb{R}^n)$ in sé. Quest'operazione si estende alle distribuzioni ponendo
$$\langle\langle \mathcal{R}_a\chi, \phi\rangle\rangle = \frac{1}{|a|^n}\langle\langle \chi, [\mathcal{R}_{1/a}\phi]\rangle\rangle\,.$$
Per capire la ragione di questa definizione, si esamini il comportamento sulle distribuzioni regolari:
$$\langle\langle \mathcal{R}_a T_f, \phi\rangle\rangle = \frac{1}{|a|^n}\langle\langle T_f, [\mathcal{R}_{1/a}\phi]\rangle\rangle = \frac{1}{|a|^n}\int_{\mathbb{R}^n}[f(x)\phi(x/a)]\,\mathrm{d}x$$
$$= \int_{\mathbb{R}^n} f(ax)\phi(x)\,\mathrm{d}x\,.$$

- il PRODOTTO CON UNA FUNZIONE. Sia $g \in C^\infty(\mathbb{R}^n)$ e sia χ una distribuzione. Il prodotto $g\chi$ si definisce ponendo:

$$\langle\langle g\chi, \phi\rangle\rangle = \langle\langle \chi, g\phi\rangle\rangle .$$

  Ciò è lecito perché abbiamo richiesto che g sia di classe C^∞.

- la DERIVATA. La formula di integrazione per parti,

$$\int_{\mathbb{R}^n} f_{x^i}(x)\phi(x)\,\mathrm{d}x = -\int_{\mathbb{R}^n} f(x)\phi_{x^i}(x)\,\mathrm{d}x$$

  suggerisce di definire

$$\langle\langle D\chi, \phi\rangle\rangle = -\langle\langle \chi, D\phi\rangle\rangle .$$

  Si vede facilmente che la trasformazione così definita è lineare e continua da $(\mathcal{D}(\mathbb{R}^n))'$ in sé. Questa trasformazione si chiama la i-ma derivata parziale della distribuzione χ.

  Più in generale, per ogni multiindice α si definisce

$$\langle\langle D^\alpha \chi\phi\rangle\rangle = (-1)^{|\alpha|}\langle\langle \chi, D^\alpha\phi\rangle\rangle .$$

Osservazione 363 Non si confonda il simbolo D^α di derivata col simbolo D^f che indica una distribuzione regolare! ∎

E' facile provare:

Teorema 364 *Ciascuna delle trasformazioni appena introdotte opera da $(\mathcal{D}(\mathbb{R}^n))'$ in sé ed è continua.*

Osservazione 365 Sia $t \to \chi(t)$ una funzione da $\mathbb{R}$ in $(\mathcal{D}(\mathbb{R}^n))'$. Si può definire la derivata di questa funzione rispetto al parametro t nel punto t_0 ponendo

$$\lim_{h\to 0} \frac{\chi(t_0+h)-\chi(t_0)}{h} ,$$

il limite essendo inteso nel senso di $(\mathcal{D}(\mathbb{R}^n))'$. Questo limite, se esiste, si indicherà col simbolo $\chi'(t_0)$ oppure col simbolo $\mathrm{d}\chi(t_0)/\,\mathrm{d}t$, da non confondere con $D\chi(t_0)$, derivata della distribuzione $\chi(t_0)$:

$$\frac{\mathrm{d}\chi(t_0)}{\mathrm{d}t} = \chi'(t_0) = \lim_{h\to 0} \frac{\chi(t_0+h)-\chi(t_0)}{h}$$

(il limite è nel senso di $(\mathcal{D}(\mathbb{R}^n))'$).

In particolare, mentre $D\chi(t_0)$ esiste sempre, $\chi'(t_0)$ può esistere o meno e, se esiste, non ha relazione con $D\chi(t_0)$. C'è però un caso in cui la relazione esiste, ed è il caso particolare in cui

$$\chi(t) = \mathcal{T}_{-t}\chi\,.$$

In tal caso,

$$\begin{aligned}
&\langle\langle \chi'(t_0), \phi\rangle\rangle = \lim_{h\to 0}\langle\langle \frac{\mathcal{T}_{-[t_0+h]}\chi - \mathcal{T}_{-t_0}\chi}{h}, \phi\rangle\rangle = \lim_{h\to 0}\langle\langle \chi, \frac{\mathcal{T}_{t_0+h}\phi - \mathcal{T}_{t_0}\phi}{h}\rangle\rangle \\
&= \lim_{h\to 0}\langle\langle \chi, \frac{\phi(t-(t_0+h)) - \phi(t-t_0)}{h}\rangle\rangle = -\langle\langle \chi, \mathcal{T}_{t_0}\phi'\rangle\rangle \\
&= -\langle\langle \mathcal{T}_{-t_0}\chi, \phi'\rangle\rangle = \langle\langle D\mathcal{T}_{-t_0}\chi, \phi\rangle\rangle\,.
\end{aligned}$$

Scegliendo $t_0 = 0$ si trova

$$D\chi = (\mathcal{T}_{-t}\chi)'_{t=0} = \lim_{h\to 0}\frac{\mathcal{T}_{-h}\chi - \chi}{h}\,,$$

eguaglianza che estende alle distribuzioni l'usuale definizione di derivata come limite del rapporto incrementale. ∎

Sottolineiamo che la derivata di una distribuzione è ancora una distribuzione e quindi è ulteriormente derivabile. Ossia, *le distribuzioni hanno derivate di ogni ordine.* Le funzioni localmente integrabili identificano distribuzioni regolari. La derivata della distribuzione D_f si chiama la DERIVATA DI f NEL SENSO DELLE DISTRIBUZIONI. Dunque, *ogni funzione localmente integrabile è derivabile nel senso delle distribuzioni.* Però, la derivata di D_f non è generalmente una distribuzione regolare. E' però facile verificare che se $f \in C^k(\mathbb{R}^n)$ e se α è un multiindice con $|\alpha| \leq k$, si ha

$$D^\alpha T_f = T_{D^\alpha f}\,.$$

Le relazioni tra la derivata in senso ordinario e nel senso delle distribuzioni sono nell'esempio seguente. Per semplicità consideriamo il caso delle funzioni su $\mathbb{R}$ ma ovviamente l'esempio si estende al caso delle derivate parziali.

Esempio 366 Sia $f(x)$ una funzione di classe $C^1(\mathbb{R})$. Sia D_f la distribuzione regolare corrsipondente ad f. La distribuzione D_f, come tutte le distribuzioni, è derivabile e, dalla definizione di derivata e di distribuzione regolare,

$$\begin{aligned}
&\langle\langle D(D_f), \phi\rangle\rangle = -\langle\langle D_f, D\phi\rangle\rangle = \int_{-\infty}^{+\infty} f(x)\phi'(x)\,\mathrm{d}x \\
&= -\int_{-\infty}^{+\infty} f'(x)\phi'(x)\,\mathrm{d}x = \langle\langle D_{f'}, \phi\rangle\rangle\,.
\end{aligned}$$

Ossia, se $f \in C^1(\mathbb{R})$ si ha

$$D(D_f) = D_{f'} \,. \tag{4.4}$$

Calcoliamo ora la derivata di alcune distribuzioni regolari che non corrispondono a funzioni derivabili.

Esempio 367 Sia $n = 1$ e sia $h(x)$ la FUNZIONE DI HEAVISIDE

$$h(x) = \begin{cases} 0 & \text{se} \quad x < 0 \\ 1 & \text{se} \quad x > 0 \,. \end{cases}$$

Questa funzione non è continua e quindi non è derivabile. E' però localmente integrabile e quindi è derivabile nel senso delle distribuzioni; ossia è derivabile la distribuzione regolare D_h. Per definizione si ha

$$\langle\!\langle D(D_h), \phi \rangle\!\rangle = -\langle\!\langle D_h, D\phi \rangle\!\rangle = -\int_0^{+\infty} \phi'(x)\, \mathrm{d}x = \phi(0) \,.$$

Dunque,

$$D(D_h) = \delta$$

non è una distribuzione regolare.

Iterando questo procedimento si vede che

$$D^n(D_h) = (-1)^{n-1} \phi^{(n-1)}(0) \,.$$

E, ovviamente, anche δ è derivabile, con

$$\langle\!\langle D^n \delta, \phi \rangle\!\rangle = -\phi^{(n)}(0) \,. \quad \blacksquare$$

Esempio 368 Sia ancora $n = 1$ e consideriamo la funzione $f(x) = \log|x|$. Questa funzione è localmente integrabile, e quindi ammette derivata nel senso delle distribuzioni. Però la sua derivata non potrà essere identificata da $1/x$ che, non essendo localmente integrabile, non definisce una distribuzione. Per calcolare la derivata di $\chi = D_f$, rifacciamoci direttamente alla definizione. E':

$$\begin{aligned}
&\langle\!\langle D\chi, \phi \rangle\!\rangle = -\langle\!\langle \chi, \phi' \rangle\!\rangle \\
&= -\lim_{\substack{a \to 0^- \\ b \to 0^+}} \left\{ \int_{-\infty}^a \phi'(x) \log|x|\, \mathrm{d}x + \int_b^{+\infty} \phi'(x) \log x\, \mathrm{d}x \right\} \\
&= -\lim_{\substack{a \to 0^- \\ b \to 0^+}} \left\{ \phi(a)\log|a| - \int_{-\infty}^a \phi(x) \frac{1}{x}\, \mathrm{d}x - \phi(b)\log b - \int_b^{+\infty} \phi(x)\frac{1}{x}\, \mathrm{d}x \right\} \,.
\end{aligned}$$

Quest'ultima espressione è la derivata di χ nel senso delle distribuzioni. Dunque, l'espressione della derivata distribuzionale di $\log|x|$ fa intervenire la funzione $1/x$, ma non è da essa rappresentata.

Per vedere un altro esempio, consideriamo invece la funzione

$$f(x)=\begin{cases}\log x & \text{se} \quad x>0\\ 0 & \text{se} \quad x<0\,.\end{cases}$$

Indicando ancora con χ la distribuzione regolare D_f, si trova:

$$\begin{aligned}&\langle\langle Df,\phi\rangle\rangle=-\langle\langle\chi,\phi'\rangle\rangle\\ &=-\lim_{a\to 0+}\int_a^{+\infty}\phi'(x)\log x\,\mathrm{d}x\\ &=-\lim_{a\to 0+}\left\{-\phi(a)\log a-\int_a^{+\infty}\phi(x)\frac{1}{x}\,\mathrm{d}x\right\}\,. \quad \blacksquare\end{aligned}$$

Studiamo ora le derivazione per serie. Lavoriamo ancora con funzioni di una variabile.

Esempio 369 Sia $\sum_{n=0}^{+\infty}\chi_n$ una serie di distribuzioni. Supponiamo che la serie converga nel senso di $(\mathcal{D}(\mathbb{R}))'$, ossia supponiamo che esista il limite

$$\chi=\lim_N\sum_{n=0}^{N}\chi_n\,.$$

Si è visto che l'operazione di derivazione, da $(\mathcal{D}(\mathbb{R}))'$ in sé è lineare e continua. Dunque essa si scambia sia con le somme finite che col segno di limite; ossia si scambia col segno di serie:

$$D\left(\sum_{n=0}^{+\infty}\chi_n\right)=\sum_{n=0}^{+\infty}D\chi_n\,.$$

Siano in particolare le χ_n distribuzioni regolari, $\chi_n=D_{f_n}$. Supponiamo che ciascuna funzione f_n sia di classe C^1 e supponiamo che la serie

$$\sum_{n=0}^{+\infty}f_n(x)$$

converga uniformemente su $\mathbb{R}$. Si vede facilmente che ciò implica la convergenza in $(\mathcal{D}(\mathbb{R}))'$ della serie delle distribuzioni regolari D_{f_n} e quindi

$$D\left(\sum_{n=0}^{+\infty}D_{f_n}\right)=\sum_{n=0}^{+\infty}D(D_{f_n})=\sum_{n=0}^{+\infty}D_{f_n'}\,.$$

Quest'uguaglianza si scrive più comunemente nella forma

$$D\left(\sum_{n=0}^{+\infty} f_n\right) = \sum_{n=0}^{+\infty} f_n' \,. \quad \blacksquare \tag{4.5}$$

4.1.3.1 Prodotto diretto di distribuzioni

Sia $\phi \in \mathcal{D}(\mathbb{R}^n)$. Suddividiamo le componenti di $x \in \mathbb{R}^n$ in due parti, rappresentando

$$x = \begin{bmatrix} y \\ z \end{bmatrix}, \qquad y \in \mathbb{R}^k\,, \quad z \in \mathbb{R}^{n-k}\,.$$

Allora, per ogni fissato z, la funzione $y \to \phi(y, z)$ appartiene a $\mathcal{D}(\mathbb{R}^k)$ e quindi per ogni z risulta definita la funzione

$$z \to \langle\langle \chi, \phi(\cdot, z)\rangle\rangle = \psi(z)\,.$$

In quest'espressione, $\chi \in (\mathcal{D}(\mathbb{R}^k))'$. Questa funzione ha supporto compatto, perché la ϕ ha supporto compatto, ed inoltre è di classe C^∞. Limitiamoci a considerare il caso in cui z ha dimensione 1 e facciamo vedere che $\psi(z)$ è di classe C^1, con

$$\frac{\mathrm{d}}{\mathrm{d}z}\psi(z) = \langle\langle \chi, \phi_z(\cdot, z)\rangle\rangle\,.$$

Ciò fatto, è ovvio che il medesimo argomento si possa ripetere per le derivate successive, e che si possa facilmente adattare al caso in cui $z \in \mathbb{R}^{n-k}$ con $n - k > 1$.

Osserviamo che, per ogni z fissato,

$$\frac{\phi(\cdot, z + h) - \phi(\cdot, z)}{h} \longrightarrow \phi_z(\cdot, z) \tag{4.6}$$

nel senso di $\mathcal{D}(\mathbb{R}^k)$. Infatti, i rapporti incrementali hanno supporto in un medesimo insieme compatto e inoltre

$$\left|\frac{\phi(\cdot, z + h) - \phi(\cdot, z)}{h} - \phi_z(\cdot, z)\right| \le h\left[\max_{(y,z)} ||\nabla\phi(y, z)||\right]\,.$$

Ciò prova la convergenza in (4.6), uniforme su $\mathbb{R}^k$. In modo analogo si prova la convergenza di tutte le derivate successive. Di conseguenza,

$$\frac{\mathrm{d}}{\mathrm{d}z}\langle\langle \chi, \phi(\cdot, z)\rangle\rangle = \lim_{h\to 0}\langle\langle \chi, \frac{\phi(\cdot, z + h) - \phi(\cdot, z)}{h}\rangle\rangle = \langle\langle \chi, (\partial\phi(\cdot, z)/\partial z)\rangle\rangle\,. \tag{4.7}$$

In particolare, la funzione $z \to \langle\langle \chi, \phi(\cdot, z)\rangle\rangle$, che ha supporto compatto, è continuamente derivabile.

Notando che anche la funzione $\phi_z(y, z)$ è una funzione test, l'argomento precedente può iterarsi e si conclude che la funzione $z \to \langle\langle \chi, \phi(\cdot, z)\rangle\rangle$ è una funzione test. Dunque, ad essa può applicarsi una $\xi \in (\mathcal{D}(\mathbb{R}^{n-k}))'$, ossia si può considerare il funzionale

$$\phi \to \langle\langle \xi \,,\, \langle\langle \chi, \phi\rangle\rangle \,\rangle\rangle$$

Mostreremo che questa funzione dipende con continuità da $\phi \in \mathcal{D}(\mathbb{R}^n)$. Si è così costruita una distribuzione di $\mathcal{D}(\mathbb{R}^n)$ a partire da una di $\mathbb{R}^k$ e di una di $\mathbb{R}^{n-k}$. Questa distribuzione si chiama il PRODOTTO DIRETTO delle due distribuzioni χ e ξ e si indica col simbolo $\chi\xi$. E' facile vedere che

$$\chi\xi = \xi\chi \,.$$

Per completare questo argomento, dobbiamo provare la continuità di $\phi \to \langle\langle \xi \,,\, \langle\langle \chi, \phi\rangle\rangle \,\rangle\rangle$ e per questo basta notare la continuità della trasformazione

$$\phi \longrightarrow \langle\langle \chi, \phi\rangle\rangle \,, \qquad \mathcal{D}(\mathbb{R}^n) \longrightarrow \mathcal{D}(\mathbb{R}^{n-k}) \,.$$

Sia per questo $\phi_n \to 0$ in $\mathcal{D}(\mathbb{R}^n)$ e proviamo che

$$\psi_n(z) = \langle\langle \chi, \phi_n\rangle\rangle \to 0 \text{in} \mathcal{D}(\mathbb{R}^{n-k}) \,.$$

Il fatto che le $\psi_n = \langle\langle \chi, \phi_n\rangle\rangle$ abbiano supporto in un medesimo compatto K è ovvio. Proviamo che $\psi_n \to 0$ uniformemente. La convergenza delle derivate delle ψ_n si potrà poi provare in modo analogo, grazie alla (4.7).

La convergenza puntuale a zero è ovvia perché $(\phi_n(\cdot, z))$ tende a zero in $\mathcal{D}(\mathbb{R}^k)$. Per assurdo, la convergenza di (ψ_n) non sia uniforme. Allora, esiste $\epsilon_0 > 0$ ed esiste una successione (z_k) tale che

$$|\psi_n(z_k)| > \epsilon_0 \,.$$

La (z_k) prende valore nel compatto K e quindi, passando ad una sottosuccessione, si può assumere $z_k \to z_0$.

Dato che $\phi_k \to 0$ nel senso di $\mathcal{D}(\mathbb{R}^n)$, si ha

$$\phi_k(\cdot, z_0) \to 0 \,, \qquad \phi_k(\cdot, z_0) - \phi_k(\cdot, z_k) \to 0 \qquad \text{in } \mathcal{D}(\mathbb{R}^k) \,.$$

Scrivendo

$$\langle\langle \chi, \phi_k(\cdot, z_0)\rangle\rangle = \langle\langle \chi, \phi_k(\cdot, z_0) - \phi_k(\cdot, z_k))\rangle\rangle + \underbrace{\langle\langle \chi, \phi_k(\cdot, z_k))\rangle\rangle}_{=\psi_k(z_k)}$$

si vede che

$$\langle\langle \chi, \phi_k(\cdot, z_0)\rangle\rangle > \frac{\epsilon_0}{2}$$

in contrasto col fatto che, invece, $\langle\langle \chi, \phi_k(\cdot, z_0)\rangle\rangle \to 0$ dato che, come abbiamo notato, $\phi_k(\cdot, z_0)$ tende a zero in $\mathcal{D}(\mathbb{R}^k)$.

Esempio 370 Sia $n = 2$ e sia

$$x = \begin{bmatrix} y \\ z \end{bmatrix} .$$

Siano $\delta(y)$ e $\delta(z)$ la delta di Dirac di $\mathbb{R}$. La δ di Dirac non è una funzione e quindi la notazione $\delta(y)$ non indica il "valore in y". Questa notazione si usa per dire che $\delta(y)$ agisce sulla variabile y. Ossia che:

$$\langle\langle \delta(y), \phi(y, z)\rangle\rangle = \phi(0, z) .$$

Analogo significato per $\delta(z)$. Sia invece $\delta(x) = \delta(y, z)$ la delta di Dirac su $\mathbb{R}^2$. Mostriamo che

$$\delta(y, z) = \delta(x)\delta(z) .$$

Questo si vede facilmente notando che

$$\psi(z) = \langle\langle \delta(y), \phi(y, z)\rangle\rangle = \phi(0, z)$$

e quindi

$$\langle\langle \delta(z), \psi(z)\rangle\rangle = \phi(0, 0) = \langle\langle \delta(x), \phi\rangle\rangle . \quad \blacksquare$$

4.1.4 Il supporto di una distribuzione

Una distribuzione non è una funzione nel senso ordinario e quindi non ha senso parlare del valore che essa assume in un punto. E' però possibile dire quando essa è nulla su un aperto Ω: diciamo che una distribuzione χ è nulla sull'aperto Ω quando

$$\langle\langle \chi, \phi\rangle\rangle = 0$$

per ogni ϕ il cui supporto è contenuto in Ω. Si chiama SUPPORTO di χ il complementare dell'unione degli aperti Ω su cui χ è nulla.

Diremo che due distribuzioni χ' e χ'' sono uguali su un aperto Ω quando $\chi' - \chi''$ è nulla su Ω.

Esempio 371 Sia χ una distribuzione a supporto in una palla $||x|| < R$ di $\mathbb{R}^n$ e sia $v \in \mathbb{R}^n$ un vettore fissato. Consideriamo le traslazioni $\mathcal{T}_{nv}$ e mostriamo che

$$\lim \mathcal{T}_{nv}\chi = 0\,.$$

Ricordando la definizione di limite in $(\mathcal{D}(\mathbb{R}^n))'$, dobbiamo far vedere che

$$\lim\langle\langle \mathcal{T}_{nv}\chi, \phi\rangle\rangle = \lim\langle\langle \chi, \mathcal{T}_{-nv}\phi\rangle\rangle = 0\,.$$

Questo è ovvio perché per ogni valore di n abbastanza grande, il supporto di $\mathcal{T}_{-nv}\phi$ non interseca la palla $||x|| < R$ e quindi, per ogni valore di n abbastanza grande,

$$\langle\langle \chi, \mathcal{T}_{-nv}\phi\rangle\rangle = 0\,. \quad \blacksquare$$

Supponiamo ora che una funzione $f(x)$ sia localmente integrabile su $\mathbb{R}$ e che essa sia derivabile (come funzione) su un aperto Ω con derivata che è localmente integrabile su Ω; ossia con derivata $f'(x)$ che è integrabile su ogni compatto contenuto in Ω.

Sia χ la *derivata della distribuzione* D_f. Mostriamo che su Ω vale $\chi = D_{f'}$. Infatti, se ϕ ha supporto in Ω, si ha

$$\langle\langle \chi, \phi\rangle\rangle = -\langle\langle D_f, \phi'\rangle\rangle = -\int_{\mathbb{R}} f(x)\phi'(x)\,\mathrm{d}x$$

per definizione di derivata nel senso delle distribuzioni e per la definizione di D_f.

D'altra parte,

$$\langle\langle D_{f'}, \phi\rangle\rangle = \int_{\mathbb{R}} f'(x)\phi(x)\,\mathrm{d}x = \int_{\Omega} f'(x)\phi(x)\,\mathrm{d}x = -\int_{\Omega} f(x)\phi'(x)\,\mathrm{d}x = \langle\langle \chi, \phi\rangle\rangle\,,$$

perché il supporto di ϕ è un compatto contenuto in Ω.

Quest'osservazione si generalizza facilmente a funzioni e distribuzioni su $\mathbb{R}^n$. Anzi, si prova facilmente:

Teorema 372 *Siano χ_1 e χ_2 due distribuzioni e sia K un aperto che contiene il supporto di $\chi_1 - \chi_2$. Allora, K contiene anche il supporto di $\chi_1' - \chi_2'$.*

Ossia, se due distribuzioni operano nello stesso modo su tutte le funzioni test che hanno supporto in un aperto Ω, nello stesso modo operano anche le loro derivate.

Esempio 373 All'esempio 368 abbiamo calcolato la derivata di $\log|x|$. Consideriamo ora la funzione $\log\sqrt{x^2+y^2}$ su $\mathbb{R}^2$. Si vede facilmente che questa funzione è localmente integrabile (si passi a coordinate polari) e quindi che identifica una distribuzione regolare. Vogliamo calcolarne il laplaciano.

Un calcolo diretto mostra che per $(x,y)\neq 0$ si ha

$$\Delta\log\sqrt{x^2+y^2}=0\,.$$

Dunque, per quanto detto sopra, il laplaciano della funzione $\log\sqrt{x^2+y^2}$ *calcolata nel senso delle distribuzioni* è nullo in ogni aperto che non contiene l'origine; ossia, il supporto del laplaciano di $\log\sqrt{x^2+y^2}$ è l'origine. Calcoliamolo ora usando direttamente la definizione. Sia $\phi(x,y)$ una funzione test. Iterando la definizione delle derivate, si vede che

$$\langle\!\langle\Delta\chi,\phi\rangle\!\rangle=\langle\!\langle\chi,\Delta\phi\rangle\!\rangle$$

e quindi, se χ è la distribuzione regolare che corrisponde a $\log\sqrt{x^2+y^2}$, va calcolato

$$\int_{\mathbb{R}^2}\left[\log\sqrt{x^2+y^2}\right][\Delta\phi(x,y)]\,\mathrm{d}x\ \mathrm{d}y=\lim_{\epsilon\to 0}\int_{\mathbb{R}^2-B_\epsilon}\left[\log\sqrt{x^2+y^2}\right][\Delta\phi(x,y)]\,\mathrm{d}x\ \mathrm{d}y\,.$$

Qui, B_ϵ è una palla di raggio ϵ intorno all'origine.

Indicando con C_ϵ la frontiera di B_ϵ e integrando per parti si trova

$$\begin{aligned}&\int_{\mathbb{R}^2-B_\epsilon}\left[\log\sqrt{x^2+y^2}\right][\Delta\phi(x,y)]\,\mathrm{d}x\ \mathrm{d}y\\&=\int_{C_\epsilon}\frac{\partial}{\partial\nu}\phi(x,y)\log\epsilon\,\mathrm{d}s-\int_{C_\epsilon}\phi(x,y)\frac{\partial}{\partial\nu}\log\sqrt{x^2+y^2}\,\mathrm{d}s\,.\end{aligned}\tag{4.8}$$

In questo calcolo si è usato il fatto che $\Delta\log\sqrt{x^2+y^2}=0$ per $x^2+y^2>\epsilon^2$ e il fatto che $\phi(x,y)$ ha supporto compatto.

La derivata $\frac{\partial}{\partial\nu}$ è la derivata "esterna" alla circonferenza C_ϵ e quindi

$$\frac{\partial}{\partial\nu}f=(\nabla f)\cdot\begin{bmatrix}-\cos\theta\\-\sin\theta\end{bmatrix}\,.$$

Si noti infatti che

$$\begin{bmatrix}-\cos\theta\\-\sin\theta\end{bmatrix}$$

è il versore normale alla circonferenza orientato verso l'origine, e quindi "esterno" rispetto alla regione $\mathbb{R}^2-B_\epsilon$.

Dunque,

$$\int_{C_\epsilon} \frac{\partial}{\partial\nu}\phi(x,y)\log\epsilon\,\mathrm{d}s$$
$$= -\int_0^{2\pi} \{\phi_x(\epsilon\cos\theta,\epsilon\sin\theta)\cos\theta + \phi_y(\epsilon\cos\theta,\epsilon\sin\theta)\sin\theta\}[\epsilon\log\epsilon]\,\mathrm{d}\theta \to 0$$

per $\epsilon \to 0$. Notiamo ora che

$$\nabla \log\sqrt{x^2+y^2} = \frac{1}{x^2+y^2}\left[\begin{array}{cc} x & y \end{array}\right].$$

Si ha quindi

$$\begin{aligned}
&-\int_{C_\epsilon} \phi(x,y)\frac{\partial}{\partial\nu}\log\sqrt{x^2+y^2}\,\mathrm{d}s = \\
&-\int_0^{2\pi} \phi(\epsilon\cos\theta,\epsilon\sin\theta)\frac{1}{\epsilon^2}\left[\begin{array}{cc} \epsilon\cos\theta & \epsilon\sin\theta \end{array}\right]\left[\begin{array}{c} -\cos\theta \\ -\sin\theta \end{array}\right](\epsilon\,\mathrm{d}\theta) \\
&= \int_0^{2\pi} \phi(\epsilon\cos\theta,\epsilon\sin\theta)\,\mathrm{d}\theta \to 2\pi\phi(0,).
\end{aligned}$$

Passando al limite per $\epsilon \to 0+$ in (4.8), si trova quindi

$$\Delta\log\sqrt{x^2+y^2} = 2\pi\delta. \quad \blacksquare$$

4.2 La trasformata di Fourier di funzioni

Sia $f(x)$ una funzione definita su $\mathbb{R}$, a valori reali o complessi. La sua TRASFORMATA DI FOURIER è la funzione della variabile *reale* ω

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega t} f(t)\, \mathrm{d}t\,. \tag{4.9}$$

Noi ci limiteremo a studiare la trasformata di Fourier di funzioni definite su $\mathbb{R}$; è però importante sapere che se $f(x)$ è definita su $\mathbb{R}^n$ allora la sua trasformata di Fourier è

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-i\xi\cdot x} f(x)\, \mathrm{d}x\,, \qquad \xi \in \mathbb{R}^n$$

(in questa formula, $\xi \cdot x$ indica il prodotto scalare di ξ e di x).

Talvolta indicheremo la trasformata di Fourier di f col simbolo $\mathcal{F}(f)$.

Si noti che $\hat{f}$ denota sia la trasformata di Fourier che la TRASFORMATA DI LAPLACE di f. La trasformata di Laplace si definisce quando $f(x) = 0$ per $x > 0$ ed è data da

$$\hat{f}(\lambda) = \int_{0}^{+\infty} e^{-\lambda t} f(t)\, \mathrm{d}t$$

(indichiamo con t la variabile di integrazione perché spesso essa rappresenta il tempo).

Il parametro λ è un numero complesso ed il dominio di $\hat{f}(\lambda)$ è l'insieme dei numeri complessi per i quali l'integrale improprio converge.

Il contesto chiarisce il significato del simbolo; si noti però che se $f(x) = 0$ per $x < 0$ e se la sua trasformata di Laplace esiste per $\Re e\, \lambda > -\epsilon$, $\epsilon > 0$, allora vale

$$\mathcal{L}(f)(i\omega) = \int_{0}^{+\infty} e^{-i\omega t} f(t)\, \mathrm{d}t = \int_{-\infty}^{+\infty} e^{-i\omega t} f(t)\, \mathrm{d}t = \mathcal{F}(f)(\omega)\,.$$

Però, in generale, *la trasformata di Fourier non ammette estensione al piano complesso.*

Prima di studiare la trasformata di Fourier, è necessario dire per quale classe di funzioni essa è definita. E' immediato notare che la definizione di trasformata di Fourier ha senso se f è integrabile (nel senso di Lebesgue) su $\mathbb{R}$: se $f \in L^1(\mathbb{R})$ l'integrale che definisce $\hat{f}(\omega)$ converge per ogni ω e anzi per ogni ω vale

$$|\hat{f}(\omega)| \leq ||f||_{L^1(\mathbb{R})}\,.$$

Dunque, la trasformazione

$$f \to \hat{f}(\omega)$$

è lineare e continua da $L^1(\mathbb{R})$ in $\mathbb{C}$, per ogni ω fissato.

Quest'osservazione suggerisce di studiare la trasformata di Fourier su $L^1(\mathbb{R})$. In realtà gli spazi $L^p(\mathbb{R})$ sono troppo piccoli per molte delle applicazioni nelle quali la trasformata di Fourier interviene. Però conviene studiare prima le proprietà della trasformata di Fourier di funzioni integrabili e poi il caso delle funzioni a quadrato integrabile. Ciò indicherà la via per estendere la trasformata a spazi molto più grandi.

4.2.1 Le proprietà della trasformata di Fourier su $L^1(\mathbb{R})$

La trasformata di Fourier è ovviamente lineare,

$$\mathcal{F}(\alpha f + \beta g) = \alpha \hat{f} + \beta \hat{g}\,.$$

Le proprietà seguenti si verificano immediatamente:

$$\begin{aligned}
&\text{se} \quad g(t) = f(t-h) \quad \text{allora} \quad \hat{g}(\omega) = e^{-i\omega h}\hat{f}(\omega)\,,\\
&\text{se } a \neq 0 \text{ e} \quad g(t) = f(at) \quad \text{allora} \quad \hat{g}(\omega) = \frac{1}{|a|}\hat{f}\left(\frac{\omega}{a}\right)\,.
\end{aligned}$$

Si confrontino con le corrispondenti proprietà della trasformata di Laplace. Inoltre:

Teorema 374 *Se $f \in L^1(\mathbb{R})$ allora la sua trasformata di Fourier è* uniformemente continua *su $\mathbb{R}$.*

<u>**Dim.**</u> E':

$$\hat{f}(\omega) - \hat{f}(\omega') = \int_{-\infty}^{+\infty} [e^{i\omega} - e^{i\omega'}] f(t)\, \mathrm{d}t\,.$$

Va provato:

$$\forall \epsilon > 0\ \exists \delta = \delta_\epsilon > 0\ :\ |\omega - \omega'| < \epsilon \implies |\hat{f}(\omega) - \hat{f}(\omega')| < \epsilon\ \forall \omega\ \ \omega' \in \mathbb{R}\,.$$

Si fissa quindi $\epsilon > 0$. Sia T_ϵ tale che

$$\int_{|t|>T_\epsilon} |f(t)|\, \mathrm{d}t < \epsilon/4\,.$$

Allora,

$$
\begin{aligned}
&|f(\omega) - f(\omega')| = \left| \int_{-\infty}^{-T_\epsilon} [e^{i\omega t} - e^{i\omega' t}] f(t)\, \mathrm{d}t \right| \\
&+ \left| \int_{-T_\epsilon}^{T_\epsilon} [e^{i\omega t} - e^{i\omega' t}] f(t)\, \mathrm{d}t \right| + \left| \int_{-T_\epsilon}^{+\infty} [e^{i\omega t} - e^{i\omega' t}] f(t)\, \mathrm{d}t \right| \\
&\le \frac{\epsilon}{2} + \int_{-T_\epsilon}^{T_\epsilon} |e^{-i\omega t} - e^{-i\omega' t}|\, |f(t)|\, \mathrm{d}t \\
&\le \frac{\epsilon}{2} + \max_{t \in [-T_\epsilon, T_\epsilon]} |e^{-i\omega t} - e^{-i\omega' t}| \int_{-\infty}^{+\infty} |f(t)|\, \mathrm{d}t\,.
\end{aligned}
$$

La funzione $s \to e^s$ è uniformemente continua sui compatti e quindi esiste $\sigma = \sigma_\epsilon > 0$ tale che se

$$
|\omega t - \omega' t| \le |\omega - \omega'| T_\epsilon < \sigma_\epsilon\,, \qquad \text{ossia se} \qquad |\omega - \omega'| < \sigma_\epsilon / T_\epsilon\,,
$$

allora

$$
|e^{i\omega t} - e^{i\omega' t}| < \frac{\epsilon}{[2 \int_{-\infty}^{+\infty} |f(t)|\, \mathrm{d}t]}\,.
$$

Dunque, per $|\omega - \omega'| < \delta_\epsilon = \sigma_\epsilon / T_\epsilon$ vale

$$
|\hat{f}(\omega) - \hat{f}(\omega')| < \epsilon\,. \quad \blacksquare
$$

In modo simile si prova:

Teorema 375 *Sia (f_n) una successione in $L^1(\mathbb{R})$ $f_n \to f_0$ nella norma di $L^1(\mathbb{R})$. Allora,*

$$
\lim \hat{f}_n(\omega) = \hat{f}_0(\omega)\,,
$$

uniformemente su $\mathbb{R}$.

L'immediata dimostrazione si omette.

Ricordiamo ora[1] che se f e g sono integrabili su $\mathbb{R}$ allora la CONVOLUZIONE

$$
h(t) = (f * g)(t) = \int_{-\infty}^{+\infty} f(t-s) g(s)\, \mathrm{d}s
$$

esiste in $L^1(\mathbb{R})$ (non è detto che debba esistere puntualmente). Vale inoltre

$$
||h||_{L^1(\mathbb{R})} \le ||f||_{L^1(\mathbb{R})} ||g||_{L^1(\mathbb{R})}\,.
$$

[1] dalla teoria dell'integrazione secondo Lebesgue.

Dunque, $\hat{h}$ esiste e, per il teorema di Fubini[2],

$$\hat{h}(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega t} \int_{-\infty}^{+\infty} f(t-s)g(s)\,\mathrm{d}s\;\mathrm{d}t = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-i\omega t} f(t-s)\,\mathrm{d}t\, g(s)\,\mathrm{d}s$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-i\omega(s+r)} f(r)\,\mathrm{d}r\, g(s)\,\mathrm{d}s = \hat{f}(\omega)\hat{g}(\omega)\,.$$

Osservazione 376 La formula precedente è importantissima per le applicazioni. Si osservi che la sua dimostrazione dipende dal fatto che la misura di Lebesgue è *invariante per traslazioni* e dal fatto che $t \to e^t$ è un *omomorfismo* del *gruppo additivo* $\mathbb{R}$ nel *gruppo moltiplicativo* dei reali positivi. ∎

L'operazione di convoluzione ha le seguenti proprietà;

- $f * (g + h) = f * g + f * h$;
- $f * g = g * f$.

Per questo si parla anche di "prodotto di convoluzione". Però si tratta di un prodotto privo di elemento neutro perchè *non esiste una funzione g tale che $f * g = f$ per ogni f.*

4.2.1.1 Il teorema di Riemann-Lebesgue

Vogliamo ora studiare

$$\lim_{|\omega|\to+\infty} \hat{f}(\omega)$$

quando f è integrabile. Consideriamo prima di tutto gli operatori di traslazione su $L^1(\mathbb{R})$. Se τ è fissato, con $\mathcal{T}_\tau$ indichiamo l'operatore da $L^1(\mathbb{R})$ in sé definito da

$$(\mathcal{T}_\tau f)(t) = f(t-\tau)$$

Ovviamente:

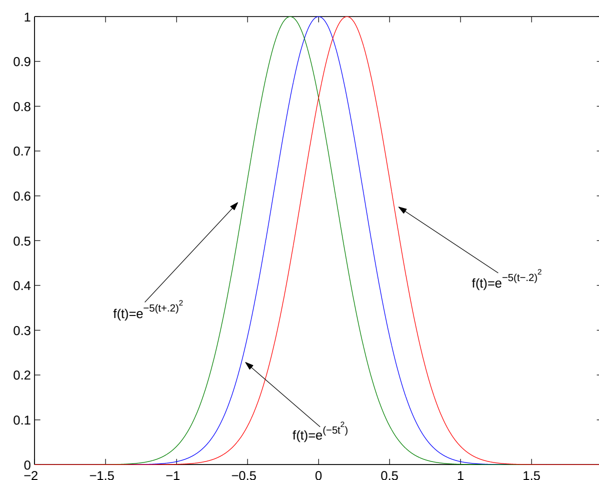
Teorema 377 *Sia τ fissato. L'operatore $\mathcal{T}_\tau$ da $L^1(\mathbb{R})$ in sé è lineare e continuo.*

Si sa, dalla teoria dell'integrazione secondo Lebesgue:

Teorema 378 (di Lebesgue) *Si fissi $f \in L^1(\mathbb{R})$ e si consideri la funzione $\tau \to \mathcal{T}_\tau f$ da $\mathbb{R}$ in $L^1(\mathbb{R})$. Questa funzione è continua.*

[2]relativo allo scambio di ordine di integrazione dell'integrale di Lebesgue.

Figura 4.1:



Usando il teorema 378 si prova:

Teorema 379 (di Riemann-Lebesgue) *Sia* $f \in L^1(\mathbb{R})$. *Vale:*

$$\lim_{|\omega|\to+\infty} \hat{f}(\omega) = 0\,.$$

Dim. Per definizione,

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega t} f(t)\,\mathrm{d}t\,. \tag{4.10}$$

Si faccia la sostituzione $t = \tau + \frac{\pi}{\omega}$ e si noti che

$$e^{-i\omega(\tau+\pi/\omega)} = e^{-i\omega\tau}e^{-i\pi} = -e^{-i\omega\tau}\,.$$

Si trova:

$$\hat{f}(\omega) = -\int_{-\infty}^{+\infty} f(\tau+\pi/\omega)e^{-i\omega\tau}\,\mathrm{d}\tau\,. \tag{4.11}$$

Sommando (4.11) e (4.10) si trova

$$\hat{f}(\omega) = \frac{1}{2}\int_{-\infty}^{+\infty} [f(\tau) - f(\tau+\pi/\omega)]e^{-i\omega\tau}\,\mathrm{d}\tau$$

così che

$$|\hat{f}(\omega)| \le \frac{1}{2}\int_{-\infty}^{+\infty} |f(\tau) - f(\tau + \pi/\omega)|\,\mathrm{d}\tau\,.$$

Essendo $f \in L^1(\mathbb{R})$, il membro destro tende a zero per $|\omega| \to +\infty$, dal teorema 378. ∎

Osserviamo ora:

Teorema 380 *Sia f una funzione derivabile su $\mathbb{R}$ e siano integrabili sia f che la sua derivata f'. Allora vale:*

$$(\mathcal{F}f')(\omega) = i\omega\hat{f}(\omega)\,.$$

Dim. L'uguaglianza

$$f(T) = f(0) + \int_0^T f'(s)\,\mathrm{d}s$$

e l'integrabilità di f e di f' mostrano

$$\lim_{|T|\to+\infty} f(T) = 0\,.$$

Scriviamo ora

$$\int_{-T}^{+T} e^{-i\omega t} f'(t)\,\mathrm{d}t = e^{-i\omega T} f(T) - e^{i\omega T} f(-T) + i\omega \int_{-T}^{+T} e^{-i\omega t} f(t)\,\mathrm{d}t\,.$$

L'asserto segue passando al limite per $|T| \to +\infty$. ∎

Nelle ipotesi del teorema 380, applicando il teorema di Riemann–Lebesgue ad f', si trova

$$\lim_{|\omega|\to+\infty} \omega\hat{f}(\omega) = 0$$

e in generale se esistono e sono integrabili f, f', $\ldots$ $f^{(k)}$, allora

$$\lim_{|\omega|\to+\infty} \omega^k\hat{f}(\omega) = 0\ : \tag{4.12}$$

la regolarità di f si riflette sul comportamento asintotico di $\hat{f}$. D'altra parte:

Teorema 381 *Se $f(t)$ e $tf(t)$ sono ambedue integrabili, allora $\hat{f}(\omega)$ è derivabile, con derivata*

$$\frac{d}{d\omega}\hat{f}(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega t}[-itf(t)]\,dt\,.$$

Dim. L'integrabilità di $tf(t)$ permette di giustificare lo scambio della derivata rispetto ad ω con l'integrale. ∎

Analogamente si vede che se f è continua e $t^k f(t)$ è integrabile allora $\hat{f}(\omega)$ è k volte derivabile. Dunque, *il comportamento asintotico di $f(t)$ si riflette sulla regolarità di $\hat{f}(\omega)$.*

E' importante ricordare queste proprietà, che sono la chiave per l'estensione della definizione della trasformata di Fourier.

4.2.2 L'antitrasformata di Fourier

Vogliamo ora capire sotto quali condizioni la conoscenza di $\hat{f}$ permette di ricostruire f. Consideriamo prima di tutto il caso particolare

$$h(t) = e^{-|t|}\,.$$

In questo caso, $\hat{h}$ si calcola facilmente usando la definizione della trasformata di Fourier,

$$\hat{h}(\omega) = \frac{2}{1+\omega^2}$$

e quindi $\hat{h}$ è integrabile.

Sia $t > 0$. Usando il metodo dei residui[3] si vede facilmente che

$$\int_{-\infty}^{+\infty} e^{i\omega t}\frac{2}{1+\omega^2}\,\mathrm{d}\omega = 2\pi i \mathrm{Res}\left[e^{izt}\frac{2}{1+z^2}, i\right] = 2\pi e^{-t}\,.$$

In modo analogo, se $t < 0$ si ha

$$\int_{-\infty}^{+\infty} e^{i\omega t}\frac{2}{1+\omega^2}\,\mathrm{d}\omega = 2\pi i \mathrm{Res}\left[e^{izt}\frac{2}{1+z^2}, -i\right] = 2\pi e^{t}\,.$$

Dunque, in quest'esempio particolare, nota $\hat{h}$, la funzione h si ritrova calcolando

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{i\omega t}\hat{h}(\omega)\,\mathrm{d}\omega\,.$$

Questa relazione tra h ed $\hat{h}$ vale molto più in generale; ma non può valere per la generica funzione integrabile perché generalmente la sua trasformata *non* è integrabile.

[3] concernente le funzioni analitiche.

Esempio 382 La funzione caratteristica dell'intervallo $[-T,T]$,

$$\chi_{[-T,T]}(t)=\begin{cases}1 & \text{se}\quad t\in[-T,T]\\ 0 & \qquad\text{altrimenti}\end{cases}$$

è integrabile. La sua trasformata di Fourier si calcola immediatamente ed è

$$\hat{f}(\omega)=2\frac{\sin\omega T}{\omega}\,.$$

Questa funzione non è integrabile secondo Lebesgue, perché si sa che non è assolutamente integrabile. ∎

Limitiamoci dunque a studiare l'invertibilità della trasformata di Fourier sotto ipotesi assai più restrittive del necessario, ma sufficiente per il seguito.

Notiamo prima di tutto che

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{2}{1+\omega^2}\,\mathrm{d}\omega=1$$

e quindi, come si è visto al paragrafo 2.4, da essa si può costruire l'identità approssimata (h_ν),

$$h_\nu(\omega)=\frac{1}{\nu}\left[\frac{1}{2\pi}\frac{2}{1+(\omega/\nu)^2}\right].$$

Vale quindi

$$f(t)=\lim_{\nu\to 0}\frac{1}{2\pi}\int_{-\infty}^{+\infty}f(t-s)\frac{1}{\nu}\frac{2}{1+(s/\nu)^2}\,\mathrm{d}s\,. \tag{4.13}$$

Inoltre, ciascuna delle funzioni h_ν è una trasformata di Fourier,

$$h_\nu(\omega)=\mathcal{F}(e^{-|\nu t|})\,. \tag{4.14}$$

Fatte queste premesse, possiamo provare:

Teorema 383 *Sia $f\in C^2(\mathbb{R})$ e siano f, f', f'' in $L^1(\mathbb{R})$. Allora vale*

$$f(t)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{+i\omega t}\hat{f}(\omega)\,d\omega\,. \tag{4.15}$$

Dim. Si è già notato che, nelle ipotesi del teorema,

$$\lim_{|\omega|\to+\infty}\omega^2\hat{f}(\omega)=0$$

(si veda (4.12)) e quindi la funzione continua $\hat{f}$ è integrabile su $\mathbb{R}$. Ciò mostra che l'integrale in (4.15) ha senso. Inoltre, l'integrabilità di f' mostra che f è limitata.

Consideriamo l'uguaglianza (4.13). Usando (4.14), questa si scrive

$$f(t) = \frac{1}{2\pi} \lim_{\nu \to 0+} \int_{-\infty}^{+\infty} f(t-s) \left[\int_{-\infty}^{+\infty} e^{-isr} e^{-\nu|r|} \, \mathrm{d}r \right] \mathrm{d}s \,.$$

La funzione $(r,t) \to f(t-s)e^{-isr}e^{-\nu|r|}$ è integrabile su $\mathbb{R}^2$ e quindi si può usare il teorema di Fubini per scambiare l'ordine di integrazione ottenendo

$$\begin{aligned} f(t) &= \lim_{\nu \to 0} \int_{-\infty}^{+\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t-s) e^{-isr} \, \mathrm{d}s \right] e^{-\nu|r|} \, \mathrm{d}r \\ &= \lim_{\nu \to 0+} \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} f(\xi) e^{-i(t-\xi)r} \, \mathrm{d}\xi \right] e^{-\nu|r|} \, \mathrm{d}r \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itr} \hat{f}(-r) \, \mathrm{d}r = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{it\omega} \hat{f}(\omega) \, \mathrm{d}\omega \,. \end{aligned}$$

Lo scambio del segno di limite col segno di integrale è lecito perché

$$\left| \left[\int_{-\infty}^{+\infty} f(\xi) e^{-i\xi r} \, \mathrm{d}\xi \right] e^{-it} e^{-\nu|t|} \right| \leq |\hat{f}(t)|$$

e, nelle ipotesi del teorema, $\hat{f}$ è integrabile.

Ciò completa la dimostrazione. ∎

La formula (4.15) si chiama la formula dell'ANTITRASFORMATA DI FOURIER. Ripetiamo che essa vale sotto condizioni assai più generali di quelle assunte nel teorema 383. Per esempio, si è visto che essa vale per la funzione $e^{-|t|}$ che non è derivabile su $\mathbb{R}$. Però l'enunciato del teorema 383 è sufficiente per il seguito.

Osservazione 384 La formula dell'antitrasformata di Fourier si può ottenere anche usando altre identità approssimate, in particolare usando quella che introduciamo in questa osservazione.

Sia

$$f(x) = e^{-x^2/2} \,. \tag{4.16}$$

Si può provare che la sua trasformata di Fourier è

$$\hat{f}(\omega) = \sqrt{2\pi} e^{-\omega^2/2} \,,$$

una funzione integrabile su $\mathbb{R}$ e non negativa. Si ricordi infatti L'INTEGRALE DI LAPLACE,

$$\int_{-\infty}^{+\infty} e^{-x^2/2}\,\mathrm{d}x = \sqrt{\pi}\,.$$

Si scriva quindi

$$\int_{-\infty}^{+\infty} e^{-x^2/2}e^{-i\omega x}\,\mathrm{d}x = e^{-\omega^2/2}\sqrt{2}\int_{-\infty}^{+\infty} e^{-[(x/\sqrt{2})+i(\omega/\sqrt{2})]^2}\,\mathrm{d}[(x/\sqrt{2})+i(\omega/\sqrt{2})]$$
$$= e^{-\omega^2/2}\sqrt{2}\int_{-\infty}^{+\infty} e^{-s^2}\,\mathrm{d}s = \sqrt{2\pi}e^{-\omega^2/2}\,.$$

Si vede immediatamente che la successione di funzioni (h_ν),

$$h_\nu(x) = \frac{1}{\nu}e^{-(x/\nu)^2/2}$$

è un'identità approssimata, la cui antitraformata di Fourier è

$$e^{-(x\nu)^2/2}\,.$$

Il teorema 383 si può provare anche usando quest'identità approssimata.

Si noti che l'identità approssimata costruita a partire da (4.16) è quella che permette di provare il Teorema di Weierstrass, Teorema 47. ∎

4.2.3 La trasformata di Fourier su $L^2(\mathbb{R})$

Abbiamo detto che intendiamo estendere la trasformata di Fourier ad un insieme assai più grande di $L^1(\mathbb{R})$. Come primo passo, estendiamola ad $L^2(\mathbb{R})$. Notiamo però che anche $L^1(\mathbb{R})\cup L^2(\mathbb{R})$ è ancora troppo piccolo per molte delle applicazioni nelle quali interviene la trasformata di Fourier.

Usiamo il simbolo[4] $\mathcal{D}(\mathbb{R})$ che indica l'insieme delle funzioni di classe C^∞ a supporto compatto in $\mathbb{R}$. Una proprietà[5] che non è difficile mostrare, è che *$\mathcal{D}(\mathbb{R})$ è denso sia in $L^1(\mathbb{R})$ che in $L^2(\mathbb{R})$.* Inoltre, ogni $f \in \mathcal{D}(\mathbb{R})$ verifica le condizioni del teorema 383 e quindi per essa vale la formula dell'antitrasformata.

Introduciamo la trasformazione lineare $\mathcal{F}$ su $\mathcal{D}(\mathbb{R})$, definita da

$$\mathcal{F}f = \hat{f}\,.$$

[4]Le proprietà di questo spazio sono state studiate al paragrafo 4.1.1. Non è però necessario conoscerle per leggere questa parte sulla trasformata di Fourier.

[5]che non abbiamo provato

Consideriamo questa come trasformazione su $L^2(\mathbb{R})$, con dominio $\mathcal{D}(\mathbb{R})$ denso in $L^2(\mathbb{R})$.

Il teorema di Riemann–Lebesgue implica che, se $f \in \mathcal{D}(\mathbb{R})$ allora

$$\lim_{\omega\to+\infty} \omega^2 \hat{f}(\omega) = 0$$

(si veda (4.12) e quindi che $\hat{f} \in L^2(\mathbb{R})$. Dunque, $\mathcal{F}$ è una trasformazione da $L^2(\mathbb{R})$ in sé, con dominio denso.

Siano ora f e g elementi di $\mathcal{D}(\mathbb{R})$. Vale:

$$\int_{-\infty}^{+\infty} f(x)g(x)\,\mathrm{d}x = \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(x)\int_{-\infty}^{+\infty} e^{+i\omega x}\hat{g}(\omega)\,\mathrm{d}\omega\ \mathrm{d}x$$
$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} f(x)e^{i\omega x}\,\mathrm{d}x\right]\hat{g}(\omega)\,\mathrm{d}\omega = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{f}(-\omega)\hat{g}(\omega)\,\mathrm{d}\omega\,.$$

Notando che

$$\hat{f}(-\omega) = \int_{-\infty}^{+\infty} f(x)e^{i\omega x}\,\mathrm{d}x = \overline{\int_{-\infty}^{+\infty}\bar{f}(x)e^{-i\omega x}\,\mathrm{d}x} = \overline{\mathcal{F}(\bar{f})(\omega)}\,,$$

l'uguaglianza precedente conduce a:

Teorema 385 *Se f, g appartengono a $\mathcal{D}(\mathbb{R})$ allora vale*

$$\int_{-\infty}^{+\infty} f(x)g(x)\,dx = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\overline{\mathcal{F}(\bar{f})}(\omega)\hat{g}(\omega)\,d\omega$$

e quindi anche

$$\int_{-\infty}^{+\infty}\bar{f}(x)g(x)\,dx = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\overline{\hat{f}(\omega)}\hat{g}(\omega)\,d\omega\,. \tag{4.17}$$

Ponendo $f = g$ in (4.20) si trova l'IDENTITÀ DI PARSEVAL

$$||f||_{L^2} = \frac{1}{\sqrt{2\pi}}||\hat{f}||_{L^2}\,.$$

Ciò mostra che

Teorema 386 *La trasformazione $\mathcal{F}$, definita su $\mathcal{D}(\mathbb{R})$, da $L^2(\mathbb{R})$ in sé è continua, con norma $1/\sqrt{2\pi}$, e quindi ammette estensione unica ad $L^2(\mathbb{R})$.*

Un calcolo analogo a quello visto sopra mostra che vale anche l'uguaglianza

$$\int_{-\infty}^{+\infty} \hat{f}(s)g(s)\,\mathrm{d}s = \int_{-\infty}^{+\infty} f(s)\hat{g}(s)\,\mathrm{d}s \tag{4.18}$$

che va sotto il nome di IDENTITÀ DI PLANCHEREL. Introducendo il simbolo

$$\langle\langle x, y\rangle\rangle = \int_{-\infty}^{+\infty} x(t)y(t)\,\mathrm{d}t = \langle x, \bar{y}\rangle\,, \tag{4.19}$$

la (4.18) si scrive

$$\langle\langle \hat{f}, g\rangle\rangle = \langle\langle f, \hat{g}\rangle\rangle\,. \tag{4.20}$$

Osservazione 387 Si noti che:

1. il funzionale $\langle\langle x, y\rangle\rangle$ è lineare sia in x, tenendo y costante, che in y, tenendo x costante. In generale si chiama BILINEARE un funzionale $F(x, y)$ definito su uno spazio lineare X tale che

$$\begin{aligned} &x \mapsto F(x, y) \text{ è continuo per ogni } y \in X \\ &y \mapsto F(x, y) \text{ è continuo per ogni } x \in X\,. \end{aligned}$$

   Si noti che in generale un funzionale bilineare non è lineare su $X \times X$, ossia non è lineare la trasformazione $(x, y) \mapsto F(x, y)$.

2. i due termini nella notazione $\langle\langle \cdot, \cdot\rangle\rangle$ si possono scambiare:

$$\langle\langle g, \hat{f}\rangle\rangle = \langle\langle \hat{f}, g\rangle\rangle = \langle\langle f, \hat{g}\rangle\rangle = \langle\langle \hat{g}, f\rangle\rangle\,. \blacksquare$$

Indichiamo con $\tilde{\mathcal{F}}$ l'estensione per continuità di $\mathcal{F}$ ad $L^2(\mathbb{R})$ (è ovvio che in pratica si userà il medesimo simbolo per le due trasformazioni). Per la trasformazione $\tilde{\mathcal{F}}$ continuano a valere le identità di Parseval e di Plancherel.

Osservazione 388 La trasformata di Fourier $\tilde{\mathcal{F}}$ è definita su $L^2(\mathbb{R})$ come estensione per continuità. In generale, l'integrale (4.9) *non* converge se la funzione è a quadrato integrabile. $\blacksquare$

Tutti gli argomenti precedenti possono ripetersi per la trasformazione $\mathcal{G}$ da $\mathcal{D} \subseteq L^2(\mathbb{R})$ in $L^2(\mathbb{R})$,

$$(\mathcal{G}\phi)(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{+i\omega t}\phi(\omega)\,\mathrm{d}\omega\,.$$

Anche $\mathcal{G}$ si estende per continuità ad $L^2(\mathbb{R})$. Provvisoriamente, indichiamo con $\tilde{\mathcal{G}}$ tale estensione. Per essa vale

$$\tilde{\mathcal{F}}\tilde{\mathcal{G}} = \tilde{\mathcal{G}}\tilde{\mathcal{F}} = I\,,$$

ossia:

Teorema 389 *La trasformazione di Fourier è biunivoca su $L^2(\mathbb{R})$ ed ha inversa continua.*

Si noti che una proprietà analoga *non* vale né su $L^1(\mathbb{R})$ né su $\mathcal{D}(\mathbb{R})$.

Ricapitolando, abbiamo esteso la trasformata e l'antitrasformata di Fourier ad $L^2(\mathbb{R})$ *per continuità.* Si sa, dal teorema 266 che l'estensione per continuità può anche costruirsi calcolando aggiunti e quindi la trasformata di Fourier si potrà estendere ad $L^2(\mathbb{R})$ calcolando $\mathcal{F}^{**}$. Conviene vedere esplicitamente anche questo metodo che suggerisce come estendere la trasformata di Fourier a spazi assai più grandi.

4.2.3.1 L'estensione "per trasposizione" della trasformata di Fourier ad $L^2(\mathbb{R})$

Sia H uno spazio di Hilbert e sia A: $H \mapsto H$ un operatore lineare continuo con dominio denso. Ricordiamo (dal paragrafo 3.6.1) che per definizione A^* è definito su tutto lo spazio H e che per definizione A^*k è l'unico elemento z tale che

$$\langle Ah, k\rangle_H = \langle h, z\rangle_H \,.$$

L'esistenza di z è garantita dal teorema di Riesz e z è unico perchè il dominio di A è denso. Inoltre, A^{**} è l'estensione continua di A ad H e quindi questo è un modo (un po' involuto) per costruire l'estensione continua di A.

Questa costruzione si può applicare in particolare allo spazio $H = L^2(\mathbb{R})$ ed all'operatore lineare continuo $\mathcal{F}$ il cui dominio $\mathcal{D}(\mathbb{R})$ è denso il $L^2(\mathbb{R})$: Si ha $\tilde{\mathcal{F}} = \mathcal{F}^{**}$. Elaborando quest'idea troveremo un modo per estendere la trasformata di Fourier a spazzi "molto grandi" che però non sono spazi di Hilbert. L'idea va rielaborata in modo da nascondere l'uso del prodotto interno e del Teorema di Riesz. Aiutano in questo l'identità di Plancherel (4.19) e il funzionale bilineare (4.18).

E' utile ricordare la notazione per i funzionali lineari su un generico spazio di Banach introdotta nel paragrafo 2.6.3.1:

$$\Psi(x) = \langle\!\langle \Psi, x\rangle\!\rangle$$

che permette di interpretare i due membri dell'identità di Plancherel: $\langle\!\langle \hat{f}, g\rangle\!\rangle$ si può interpretare come *il valore preso su g dell'elemento $\hat{f}$ dello spazio duale.*

Tornando all'identità di Plancherel, consideriamo la seguente catena di uguaglianze:

$$\begin{array}{ccccccc}
\langle\!\langle \mathcal{F}f, g\rangle\!\rangle & = & \langle \mathcal{F}f, \bar{g}\rangle & = & \langle f, \mathcal{F}^*\bar{g}\rangle & = & \langle \mathcal{F}^{**}f, \bar{g}\rangle \\
\| & & & & & & \| \\
{\scriptstyle (\text{Plancherel})} & & & & & & \\
\langle\!\langle f, \mathcal{F}g\rangle\!\rangle & & & & \langle\!\langle \tilde{\mathcal{F}}f, g\rangle\!\rangle & = & \langle\!\langle \mathcal{F}^{**}f, g\rangle\!\rangle
\end{array}$$

Dunque, possiamo intendere $\mathcal{F}^{**}f = \tilde{\mathcal{F}}f$ come *elemento dello spazio duale: quell'elemento che a $g \in \mathcal{D}(\mathbb{R})$ fa corrispondere il numero $\langle\langle f, \mathcal{F}g\rangle\rangle$.* Ossia, $\tilde{\mathcal{F}}f$ è definita mediante l'uguaglianza

$$\langle\langle \tilde{\mathcal{F}}f, g\rangle\rangle = \langle\langle f, \hat{g}\rangle\rangle \qquad \forall g \in \mathcal{D}\,. \tag{4.21}$$

Questa uguaglianza non fa più uso esplicito del Teorema di Riesz e suggerisce un modo per estendere la trasformata di Fourier ad uno spazio *molto grande:* prima si identifica uno spazio *molto piccolo,* con una topologia *molto forte,* su cui la trasformata di Fourier è continua e biunivoca. Si usa quindi un metodo "di trasposizione" per estenderla al duale dello spazio. L'idea intuitiva è che se lo spazio è "piccolo" il suo duale sarà "grande". Va notato subito però che lo spazio "piccolo" da usare non è $\mathcal{D}(\mathbb{R})$ perché la trasformata di Fourier di una $\phi \in \mathcal{D}(\mathbb{R})$ non appartiene a $\mathcal{D}(\mathbb{R})$. Introdurremo nel prossimo paragrafo lo spazio $\mathcal{S}$, più grande di $\mathcal{D}(\mathbb{R})$, ancora denso sia in $L^1(\mathbb{R})$ che in $L^2(\mathbb{R})$ e che può usarsi a questo scopo.

Osservazione 390 Concludiamo con queste osservazioni:

1. Abbiamo fatto intervenire il funzionale $\langle\langle\cdot,\cdot\rangle\rangle$, lineare nei due argomenti, invece di usare il prodotto interno. La ragione è che lo spazio $\mathcal{S}$ che introdurremo al prossimo paragrafo non è di Hilbert e quindi non ha prodotto interno.

2. E' immediato calcolare che se $\phi \in \mathcal{D}(\mathbb{R})$ si ha

$$\mathcal{F}^*\phi = \int_{-\infty}^{+\infty} e^{+i\omega t}\phi(\omega)\,\mathrm{d}\omega = 2\pi\mathcal{F}^{-1}\phi\,,$$
$$\overline{\mathcal{F}^*\phi} = \int_{-\infty}^{+\infty} e^{-i\omega t}\bar{\phi}(\omega)\,\mathrm{d}\omega = \mathcal{F}\bar{\phi}\,.$$

3. La trasformata di Fourier su $L^2(\mathbb{R})$ permette di ottenere delle formule interessanti. Presentiamo un esempio. Applichiamo l'identità di Parseval alla funzione $\chi_{[-1,1]}$, studiata all'esempio 382. La sua trasformata di Fourier

$$2\frac{\sin\omega}{\omega}$$

è in $L^2(\mathbb{R})$, come si verifica immediatamente, e come deve essere perché $\chi_{[-1,1]}$ è a quadrato integrabile. Dunque,

$$\int_{-\infty}^{+\infty}\left[2\frac{\sin\omega}{\omega}\right]^2\mathrm{d}\omega = 2\pi\int_{-\infty}^{+\infty}[\chi_{[-1,1]}(x)]^2\,\mathrm{d}x = 4\pi\,.$$

Si trova quindi

$$\int_{-\infty}^{+\infty} \left[\frac{\sin\omega}{\omega}\right]^2 \mathrm{d}\omega = \pi\,. \quad \blacksquare$$

4.3 Distribuzioni temperate

4.3.1 Lo spazio $\mathcal{S}$ e il suo duale

L'idea per la scelta dello spazio $\mathcal{S}$ su cui definire la trasformata di Fourier è fornita dal teorema di Riemann-Lebesgue, e dalle sue conseguenze: *la regolarità* di f si trasferisce nel *comportamento asintotico* di $\hat{f}$; il *comportamento asintotico* di f si trasferisce nella *regolarità* di $\hat{f}$, e viceversa. Ciò suggerisce di introdurre lo spazio[6] $\mathcal{S}$ i cui elementi sono le funzioni $\phi \in C^\infty(\mathbb{R})$ tali che:

$$\lim_{|x|\to+\infty} x^k \phi^{(n)}(x) = 0 \qquad \forall k\,,\ n\,. \tag{4.22}$$

E' chiaro che $\mathcal{S}$ è un sottoinsieme sia di $L^1(\mathbb{R})$ che di $L^2(\mathbb{R})$ e che per gli elementi di $\mathcal{S}$ valgono sia la formula della trasformata che dell'antitrasformata di Fourier:

$$\hat{\phi}(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega t}\phi(t)\,\mathrm{d}t\,, \qquad \phi(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{i\omega t}\hat{\phi}(\omega)\,\mathrm{d}\omega\,.$$

Gli integrali convergono per ogni ω.

Vogliamo *mimare* su $\mathcal{S}$ la costruzione della trasformata ottenuta *per trasposizione* su $L^2(\mathbb{R})$. Per questo è necessario munire $\mathcal{S}$ di una topologia la quale tenga conto della proprietà (4.22). E' un fatto che ciò non può farsi introducendo una norma in $\mathcal{S}$. D'altra parte, la definizione della topologia porterebbe via troppo tempo. Dunque limitiamoci a introdurre un concetto di convergenza di successioni in $\mathcal{S}$.

Per definizione,

$$\lim \phi_n(x) = 0 \quad \text{in } \mathcal{S}$$

quando *per ogni* k intero non negativo e *per ogni* r intero non negativo si ha

$$\lim_n x^k \phi_n^{(r)}(x) = 0$$

uniformemente su $\mathbb{R}$.

[6] dovremmo usare la notazione più completa $\mathcal{S}(\mathbb{R})$ perché una definizione analoga a quella che andiamo a dare può ripetersi anche per funzioni su $\mathbb{R}^n$, ottenendo lo spazio $\mathcal{S}(\mathbb{R}^n)$. Però in questa parte lavoreremo solamente con funzioni di una variabile e quindi useremo la notazione più semplice $\mathcal{S}$.

Esplicitamente questo vuol dire:

$$\forall \epsilon > 0 \; \exists N = N(\epsilon, k, r) : n > N(\epsilon, k, r) \quad \Longrightarrow \quad |x^k \phi_n^{(r)}(x)| < \epsilon \,.$$

Si è scritta esplicitamente questa condizione per *sottolineare che la convergenza NON è uniforme in* k *ed* r.

Definiamo inoltre:

$$\lim \phi_n = \phi_0 \qquad \Longleftrightarrow \qquad \lim [\phi_n - \phi_0] = 0 \,.$$

Lo spazio lineare $\mathcal{S}$, dotato della definizione di convergenza appena introdotta, si chiama lo spazio delle FUNZIONI RAPIDAMENTE DECRESCENTI SU $\mathbb{R}$.

Sia ora A un funzionale su $\mathcal{S}$ oppure una trasformazione da $\mathcal{S}$ in sé. Diremo che A è continuo quando

$$\lim \phi_n = \phi_0 \quad \Longrightarrow \quad \lim A\phi_n = A\phi_0 \,.$$

Mostriamo alcuni esempi di trasformazioni continue da $\mathcal{S}$ in sé.

Esempio 391 le trasformazioni

$$\phi \to \phi + \psi \,, \qquad \phi \to \alpha\phi$$

(con $\psi \in \mathcal{S}$ fissata) sono continue e la seconda è anche lineare.

Più ancora, sia p un polinomio. La trasformazione lineare

$$\phi(x) \to p(x)\phi(x)$$

è continua.

E' anche continua le trasformazione

$$\phi(x) \to \phi(ax) \text{ se } a \neq 0$$

Invece, $\phi(0x) = \phi(0) \notin \mathcal{S}$ a meno che sia $\phi(0) = 0$).

Per ogni ω fissato, la trasformazione

$$\phi(x) \to e^{i\omega x}\phi(x)$$

è continua. In generale, se $f \in C^\infty(\mathbb{R})$ e se f e tutte le sue derivate crescono in modo polinomiale, la trasformazione lineare

$$\phi(x) \to f(x)\phi(x)$$

è continua. ∎

Inoltre:

Teorema 392 *Si ha:*

1. *L'operazione di derivazione di qualunque ordine k:*
$$\phi \to D^k\phi$$
è continua da $\mathcal{S}$ in sé;

2. *La trasformata di Fourier $\mathcal{F}$ e la antitrasformata trasformano $\mathcal{S}$ in sé, sono continue e biunivoche.*

In particolare, la trasformata di Fourier è continua surgettiva ed invertibile con inversa continua da $\mathcal{S}$ in sé.

Dim. La linearità delle trasformazioni è ovvia ed è facile vedere che le derivate trasformano $\mathcal{S}$ in sé e sono trasformazioni continue. Esaminiamo la trasformata di Fourier.

Mostriamo prima che $\mathcal{F}$ trasforma $\mathcal{S}$ in sé. Sia per questo $\phi \in \mathcal{S}$ e sia $\hat{\phi}$ la sua trasformata:
$$\hat{\phi}(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega x}\phi(x)\,\mathrm{d}x$$

La funzione $x^k\phi(x)$ è in $\mathcal{S}$ per ogni k e quindi
$$\lim_{|\omega|\to+\infty} D^k\hat{\phi}(\omega) = (-i)^k \lim_{|\omega|\to+\infty}\int_{-\infty}^{+\infty} e^{-i\omega x}(x^k\phi(x))\,\mathrm{d}x = 0$$
Per il Teorema di Riemann-Lebesgue.

Analogamente,
$$\lim_{|\omega|\to+\infty}\omega^n\hat{\phi}(\omega) = \lim_{|\omega|\to+\infty}\int_{-\infty}^{+\infty}\left[\omega^n e^{-i\omega x}\right]\phi(x)\,\mathrm{d}x$$
$$= \lim_{|\omega|\to+\infty}\int_{-\infty}^{+\infty}\left[(i)^n\frac{\mathrm{d}^n}{\mathrm{d}x^n}e^{-i\omega x}\right]\phi(x)\,\mathrm{d}x = (-i)^n\int_{-\infty}^{+\infty} e^{-i\omega x}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\phi(x)\,\mathrm{d}x = 0\,.$$

Naturalmente i due casi si possono combinare ottenendo
$$\lim_{|\omega|\to+\infty}\omega^n\frac{\mathrm{d}^k}{\mathrm{d}\omega^k}\hat{\phi}(\omega) = 0\,.$$
Dunque $\mathcal{F}$ trasforma $\mathcal{S}$ in sé. Calcoli analoghi mostrano che anche la antitrasformata di Fourier trasforma $\mathcal{S}$ in sé.

Proviamo la continuità. Sia (ψ_n) una qualsiasi successione tendente a zero in $\mathcal{S}$. Proviamo che la successione $(\mathcal{F}\psi_n)$ tende a zero. Fissiamo per questo k ed r e consideriamo

$$\begin{aligned}\omega^k \frac{\mathrm{d}^r}{\mathrm{d}\omega^r}\int_{-\infty}^{+\infty} e^{-i\omega x}\psi_n(x)\,\mathrm{d}x &= \omega^k \int_{-\infty}^{+\infty}(-i)^r e^{-i\omega x}[x^r\psi_n(x)]\,\mathrm{d}x\\ &= \int_{-\infty}^{+\infty}(-i)^r(-i)^{-k}\left[\frac{\mathrm{d}^k}{\mathrm{d}x^k}e^{-i\omega x}\right][x^r\psi_n(x)]\,\mathrm{d}x\,. \end{aligned}\qquad (4.23)$$

Integriamo per parti tenendo conto che

$$\lim_{|x|\to+\infty}[x^r\psi_n(x)] = 0\,.$$

Si trova:

$$\begin{aligned}(4.23) &= (-i)^{r-k}(-1)^k\int_{-\infty}^{+\infty} e^{-i\omega x}\left[\frac{\mathrm{d}^k}{\mathrm{d}x^k}x^r\psi_n(x)\right]\,\mathrm{d}x\\ &= (-i)^{r-k}(-1)^k\int_{-\infty}^{+\infty}\frac{e^{-i\omega x}}{1+x^2}\left[(1+x^2)\frac{\mathrm{d}^k}{\mathrm{d}x^k}x^r\psi_n(x)\right]\,\mathrm{d}x\,.\end{aligned}$$

Sia ora $\epsilon > 0$. Esiste $N(\epsilon, k, r)$ tale che

$$n > N(\epsilon, k, r) \quad\Longrightarrow\quad \left|(1+x^2)\frac{\mathrm{d}^k}{\mathrm{d}x^k}x^r\psi_n(x)\right| < \epsilon$$

così che, per tali indici n si ha anche

$$\left|\omega^k\frac{\mathrm{d}^r}{\mathrm{d}\omega^r}\int_{-\infty}^{+\infty} e^{-i\omega x}\psi_n(x)\,\mathrm{d}x\right| \le \pi\epsilon\,.$$

Questo prova la continuità di $\mathcal{F}$. La trasformazioine $\mathcal{F}^{-1}$ si studia in modo analogo.

La trasformazione $\mathcal{F}$ è suriettiva perché ammette l'inversa. ∎

Con $\mathcal{S}'$ indichiamo lo spazio lineare dei funzionali lineari e continui su $\mathcal{S}$, dotato della relazione di convergenza seguente

$$\lim l_n = l_0 \quad\Longleftrightarrow\quad \lim l_n(\phi) = l_0(\phi) \qquad \forall\phi\in\mathcal{S}\,.$$

Si confronti questa definizione di convergenza con la convergenza debole stella e con la convergenza in $(\mathcal{D}(\mathbb{R}))'$.

Gli elementi di $\mathcal{S}'$ si chiamano DISTRIBUZIONI TEMPERATE.

Come al solito, per indicare l'azione di $l \in \mathcal{S}'$ su $\phi \in \mathcal{S}$, invece di scrivere $l(\phi)$ scriveremo

$$\langle\langle l, \phi\rangle\rangle\,.$$

Mostriamo ora alcuni esempi di distribuzioni temperate.

Esempio 393 Sia l tale che

$$\langle\langle l, \phi \rangle\rangle = \phi(0)\,.$$

E' immediato verificare che l è lineare e continuo, ossia che $l \in \mathcal{S}'$.

Questa distribuzione è particolarmente importante per le applicazioni ed ha un simbolo standard: si indica col simbolo δ,

$$\langle\langle \delta, \phi \rangle\rangle = \phi(0)$$

e si chiama DELTA DI DIRAC.

E ancora immediato verificare la continuità di

$$\phi \to \sum_{k=0}^{\nu} a_k \phi^{(k)}(x_k)\,. \quad \blacksquare$$

In seguito chiariremo le relazioni tra le due distribuzioni temperate introdotte nell'esempio precedente.

Consideriamo ora:

Esempio 394 Sia $f \in L^p(\mathbb{R})$, $1 \le p \le +\infty$. E' una distribuzione temperata quella definita da

$$\phi \to \int_{-\infty}^{+\infty} f(s)\phi(s)\,\mathrm{d}s\,.$$

Questo si verifica immediatamente usando il *teorema della convergenza* dominata, se $p = 1$. Se $p > 1$ si usa la disuguaglianza di Hoelder per notare che

$$\frac{1}{x^2+1} f(x) \in L^1(\mathbb{R})\,.$$

Si scrive quindi

$$\int_{-\infty}^{+\infty} f(s)\phi(s)\,\mathrm{d}s = \int_{-\infty}^{+\infty} \frac{f(x)}{1+x^2}\left[(1+x^2)\phi(x)\right]\,\mathrm{d}x\,.$$

Sappiamo già che la trasformazione che a ϕ associa $(1+x^2)\phi(x)$ è continua. E quindi la trasformazione che stiamo studiando è continua, essendo composizione delle due trasformazioni lineari e continue

$$\phi \to (1+x^2)\phi(x)\,, \qquad \psi \to \int_{-\infty}^{+\infty} \frac{f(x)}{1+x^2}\psi(x)\,\mathrm{d}x\,. \quad \blacksquare$$

Le particolari distribuzioni introdotte nell'esempio 394 si chiamano DISTRIBUZIONI REGOLARI. Esse si indicano con un simbolo del tipo D_f oppure più spesso semplicemente f o $\int f\phi$. In analogia con questo simbolo, specialmente nei testi più applicativi, una distribuzione l si indica col simbolo

$$\phi \to \int l\phi$$

attribuendo al simbolo "$\int$" il significato del simbolo "$\langle\!\langle \cdot, \cdot \rangle\!\rangle$".

In pratica, non si distingue tra le funzioni e le distribuzioni regolari ad esse associate. Però, per maggior chiarezza in questa parte useremo il simbolo D_f per indicare la distribuzione regolare identificata dalla funzione f. Si faccia solo attenzione a non confondere D_f con Df che indica la derivata di f.

Riportiamo di nuovo l'esempio seguente:

Esempio 395 Sia (h_n) un'identità approssimata. Per ogni $\phi \in \mathcal{S}$ si ha

$$\lim_n \int_{-\infty}^{+\infty} h_n(t-s)\phi(s)\,\mathrm{d}s = \phi(t)\,,$$

si veda il paragrafo 2.4. Ciò vale in particolare per $t = 0$ e quindi

$$\lim_n \int_{-\infty}^{+\infty} h_n(-s)\phi(s)\,\mathrm{d}s = \phi(0)\,, \qquad \phi \in \mathcal{S}\,.$$

Dunque, la successione di distribuzioni regolari definite da

$$\phi \to \int_{-\infty}^{+\infty} h_n(-s)\phi(s)\,\mathrm{d}s$$

converge in $\mathcal{S}'$ alla δ di Dirac. Si dice più brevemente che "le identità approssimate approssimano la δ di Dirac". ∎

4.3.2 Distribuzioni e distribuzioni temperate

> La lettura di questo paragrafo richiede gli argomenti del paragrafo 4.1. Chi non ha studiato il paragrafo 4.1 può omettere la lettura di questo e passare ai successivi.

Scegliendo di chiamare "distribuzioni temperate" gli elementi di $\mathcal{S}'$, abbiamo suggerito l'inclusione

$$\mathcal{S}' \subseteq (\mathcal{D}(\mathbb{R}))'\,.$$

Quest'inclusione si prova facimente perché ogni $\phi \in \mathcal{D}(\mathbb{R})$ è anche in $\mathcal{S}$ e se (ϕ_n) è una successione tendente a zero nel senso di $\mathcal{D}(\mathbb{R})$, essa tende a zero anche nel senso di $\mathcal{S}$. Infatti,

$$\frac{\mathrm{d}^r}{\mathrm{d}x^r} x^k \phi_n(x)$$

è somma finita di termini del tipo

$$x^j \phi_n^{(m)}(x)$$

e queste successioni tendono a zero per $n \to +\infty$ uniformemente su $\mathbb{R}$, *perché se $\phi_n \to 0$ in $\mathcal{D}(\mathbb{R})$ allora esiste un compatto K che contiene i supporti di tutte le ϕ_n.*

Ossia, ogni elemento di $\mathcal{S}'$ ammette restrizione a $\mathcal{D}(\mathbb{R})$ e quindi identifica un funzionale lineare continuo su $\mathcal{D}$ definendo $\langle\langle l, \phi \rangle\rangle_{\mathcal{D}',\mathcal{D}}$ mediante l'uguaglianza seguente:

$$\langle\langle l, \phi \rangle\rangle_{\mathcal{D}',\mathcal{D}} =_{\mathrm{Def.}} \langle\langle l, \phi \rangle\rangle_{\mathcal{S}',\mathcal{S}} .$$

Mostriamo che, viceversa, ogni elemento di $\mathcal{D}'(\mathbb{R})$ è restrizione di *al più* un elemento di $\mathcal{S}'$. Per questo basta provare

$$\langle\langle l, \phi \rangle\rangle = 0 \qquad \forall \phi \in \mathcal{D}(\mathbb{R}) \implies l = 0 . \tag{4.24}$$

L'uguaglianza $l = 0$ va intesa in $\mathcal{S}'$, ossia va provato che $\langle\langle l, \psi \rangle\rangle = 0$ *per ogni* $\psi \in \mathcal{S}$. Questo fatto discende dal il teorema seguente, che non proviamo:

Teorema 396 *Lo spazio $\mathcal{D}(\mathbb{R})$ è denso in $\mathcal{S}$.*

Usando questo risultato proviamo l'implicazione (4.24) ossia proviamo che se $\langle\langle l, \phi \rangle\rangle = 0$ *per ogni* $\phi \in \mathcal{D}(\mathbb{R})$ allora si ha $\langle\langle l, \psi \rangle\rangle = 0$ anche *per ogni* $\psi \in \mathcal{S}$. Si noti che $\langle\langle l, \phi \rangle\rangle$ indica il valore di $l \in \mathcal{S}'$ su ϕ e che $\phi \in \mathcal{S}$ perché $\phi \in \mathcal{D}(\mathbb{R}) \subseteq \mathcal{S}$.

Sia $\psi \in \mathcal{S}$. Il Teorema 396 mostra che esiste una successione (ϕ_n) di elementi di $\mathcal{D}(\mathbb{R})$ tale che $\phi_n \to \psi$ in $\mathcal{S}$. Quindi,

$$\langle\langle l, \psi \rangle\rangle = \lim \langle\langle l, \phi_n \rangle\rangle \to 0 .$$

Essendo $\psi \in \mathcal{S}$ arbitrario segue che $l = 0$.

Ricapitolando, abbiamo visto che ogni distribuzione temperata è una distribuzione. Mostriamo però che esistono distribuzioni che *non* sono distribuzioni temperate:

Esempio 397 Consideriamo la distribuzione regolare, elemento di $\mathcal{D}'$, identificata dalla funzione $f(x) = e^{x^4}$. Essendo questa funzione localmente integrabile essa definisce una distribuzione. *Mostriamo che non definisce una distribuzione temperata.* Per vederlo, consideriamo la funzione $\psi(x) = e^{-x^2}$, che è un elemento di $\mathcal{S}$. Ovviamente,

$$\int_{-\infty}^{+\infty} e^{x^4} e^{-x^2}\,\mathrm{d}x = +\infty\,.$$

Questo mostra che e^{x^4} non definisce una distribuzione temperata regolare. Per mostrare che non è nemmeno una distribuzione temperata, usiamo di nuovo il Teorema 396. Sia $\{\phi_n\}$ una successione in $\mathcal{D}(\mathbb{R})$ che tende a ψ nel senso di $\mathcal{S}$. Possiamo supporre che per ogni x sia

$$M < \phi_1(x) \le \phi_n(x) \le \phi_{n+1}(x) \le \psi(x)\,.$$

La funzione $f(x) = e^{x^4}$ definisce quella distribuzione che associa a ϕ_n il valore

$$\langle\!\langle D_f, \phi_n \rangle\!\rangle = \int_{-\infty}^{+\infty} f(x)\phi_n(x)\,\mathrm{d}x \to \int_{-\infty}^{+\infty} f(x)\phi(x)\,\mathrm{d}x = +\infty\,.$$

per il teorema di Beppo Levi[7] e questo non potrebbe accadere se la distribuzione D_f fosse anche una distribuzione temperata. ∎

4.3.3 La trasformata di Fourier su $\mathcal{S}'$

Si è già detto che su $\mathcal{S}$ la trasformata di Fourier è definita dall'integrale (4.9), è lineare, continua iniettiva e suriettiva e che vale la formula dell'antitrasformata, che a sua volta è lineare, continua iniettiva e suriettiva.

Con queste informazioni possiamo estendere la trasformata di Fourier ad $\mathcal{S}'$.

Si ricordi che se $f \in L^2(\mathbb{R})$ la sua trasformata di Fourier può definirsi usando la formula (4.21) e che questa può interpretrarsi dicendo che $\hat{f}$ è quel funzionale lineare e continuo che a ϕ associa $\langle\!\langle f, \hat{\phi} \rangle\!\rangle$.

Traendo ispirazione da quest'osservazione, definiamo la TRASFORMATA DI FOURIER DI DISTRIBUZIONI TEMPERATE come segue: Sia $l \in \mathcal{S}'$. La sua trasformata di Fourier $\hat{l}$ è il funzionale lineare

$$\phi \to \langle\!\langle l, \hat{\phi} \rangle\!\rangle\,.$$

[7]concernente lo scmbio dell'integrale di Lebesgue con i limiti.

E' immediato verificare[8] che $\hat{l} \in \mathcal{S}'$. Quindi l'uguaglianza

$$\langle\langle \hat{l}, \phi \rangle\rangle = \langle\langle l, \hat{\phi} \rangle\rangle$$

che ora definisce la trasformata di Fourier $\hat{l}$, estende la formula di Plancherel[9]

Così come si estende ad $\mathcal{S}'$ la trasformazione $\mathcal{F}$, si estende anche l'antitrasformata $\mathcal{G}$:

$$\mathcal{G} \ : \quad \langle\langle \mathcal{G}l, \phi \rangle\rangle = \langle\langle l, \mathcal{G}\phi \rangle\rangle$$

e la relazione

$$\langle\langle \mathcal{F}\left[\mathcal{G}l\right], \phi \rangle\rangle = \langle\langle l, \mathcal{G}\left[\mathcal{F}\right]\phi \rangle\rangle = \langle\langle l, \phi \rangle\rangle$$

mostra che l'estensione di $\mathcal{G}$ è inversa destra di $\mathcal{F}$. Procedendo in modo analogo si vede che è anche inversa sinistra, e quindi che è l'antitrasformata di Fourier.

Notiamo che:

Teorema 398 *La trasformata di Fourier è lineare continua e biunivoca da $\mathcal{S}'$ in sé, con inversa continua.*

Dim. Sia

$$\lim l_n = l_0$$

ossia

$$\lim \langle\langle l_n, \phi \rangle\rangle = \langle\langle l_0, \phi \rangle\rangle \qquad \forall \phi \in \mathcal{S}\,. \tag{4.25}$$

Vogliamo provare che $\lim \hat{l}_n = \hat{l}_0$. Ogni $\hat{\phi}$ è in $\mathcal{S}$ e quindi scegliendo per ϕ la $\hat{\phi}$ la (4.25) si scrive:

$$\lim \langle\langle l_n, \hat{\phi} \rangle\rangle = \langle\langle l_0, \hat{\phi} \rangle\rangle \qquad \text{ossia} \qquad \lim \langle\langle \hat{l}_n, \phi \rangle\rangle = \langle \hat{l}_0, \phi \rangle\rangle\,.$$

Questa uguaglianza vale per ogni $\hat{\phi} \in \mathcal{S}$ ossia per ogni $\phi \in \mathcal{S}$ e quindi

$$\lim \hat{l}_n = \hat{l}_0\,,$$

ossia la proprietà che volevamo provare.

La biunivocità della trasformata si ottiene perché, come si è già notato, anche l'antitrasformata si estende ad $\mathcal{S}'$.

Infine, l'argomento precedente può usarsi anche per provare la continuità dell'antitrasformata. ∎

---

[8]notando che $\phi \to \langle\langle l, \hat{\phi} \rangle\rangle$ è un funzionale lineare e continuo, come composizione di trasformazioni lineari e continue.

[9]A questo punto non è più necessario distinguere tra $\mathcal{F}$ e la sua estensione, che verrà ancora indicata col simbolo $\mathcal{F}$. Analogamente, l'estensione di $\mathcal{G}$ che ora introduciamo sarà indicata con $\mathcal{G}$.

Osservazione 399 Ricordiamo il punto 1 dell'osservazione 390. Si capisce ora perché l'estensione per trasposizione della trasformata di Fourier ad $L^2(\mathbb{R})$ è stata fatta usando il funzionale bilineare $\langle\langle \cdot, \cdot \rangle\rangle$ invece del prodotto interno. ∎

Esempio 400 Consideriamo ora alcuni esempi.

1. Sia $f \in L^1(\mathbb{R})$ e D_f la distribuzione regolare

$$\langle\langle D_f, \phi \rangle\rangle = \int_{-\infty}^{+\infty} f(x)\phi(x)\, \mathrm{d}x\,.$$

   La trasformata di Fourier di D_f è la distribuzione

$$\phi \to \langle\langle D_f, \hat{\phi} \rangle\rangle = \int_{-\infty}^{+\infty} f(\omega) \int_{-\infty}^{+\infty} e^{-i\omega x}\phi(x)\, \mathrm{d}x\; \mathrm{d}\omega$$
$$= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} e^{-i\omega x} f(\omega)\, \mathrm{d}\omega \right] \phi(x)\, \mathrm{d}x = \langle\langle D_{\hat{f}}, \phi \rangle\rangle \;:$$

   la trasformata di Fourier della distribuzione regolare identificata dalla funzione f è la distribuzione regolare identificata dalla funzione $\hat{f}$.

2. Consideriamo ora la funzione $f(t) = e^{i\alpha t}$, che non è né integrabile né a quadrato integrabile, e quindi non ha trasformata di Fourier nel senso che abbiamo introdotto per le funzioni. Nel senso delle distribuzioni, la sua trasformata di Fourier è

$$\phi \to \langle\langle D_f, \hat{\phi} \rangle\rangle = \int_{-\infty}^{+\infty} e^{i\alpha x}\hat{\phi}(x)\, \mathrm{d}x = 2\pi\phi(\alpha)\,.$$

   Indicando con δ_α la distribuzione

$$\langle\langle \delta_\alpha, \phi \rangle\rangle = \phi(\alpha)\,,$$

   si vede che

$$\mathcal{F}(e^{i\alpha x}) = 2\pi\delta_\alpha\,.$$

   Usando le formule di Eulero,

$$\mathcal{F}(\sin \alpha x) = -i\pi\{\delta_\alpha - \delta_{-\alpha}\}\,, \qquad \mathcal{F}(\cos \alpha x) = \pi\{\delta_\alpha + \delta_{-\alpha}\}\,. \;\blacksquare$$

   Quest'esempio mostra che la trasformata di una distribuzione regolare può non essere una distribuzione regolare.

3. Sia
$$f(t) = \begin{cases} 1 & \text{se} \quad |t| < T \\ 0 & \text{se} \quad |t| > T\,. \end{cases}$$
Come si è visto all'esempio 382,
$$\hat{f}(\omega) = \frac{\sin \omega T}{\omega}\,.$$
La funzione $\hat{f}$ non appartiene ad $L^1(\mathbb{R})$; è però limitata e quindi identifica una distribuzione regolare $D_{\hat{f}}$.

   Ad $\hat{f}$ *come funzione* la formula dell'antitrasformata che abbiamo provato al paragrafo 4.2.2 *non può applicarsi.* Invece, ad essa può applicarsi la formula dell'antitrasformata nel senso che abbiamo introdotto in $\mathcal{S}'$.

4. La trasformata $\hat{\delta}$ della delta di Dirac è la distribuzione
$$\phi \to \langle\langle \delta, \hat{\phi} \rangle\rangle = \hat{\phi}(0) = \int_{-\infty}^{+\infty} \phi(s)\,\mathrm{d}s\,.$$
E' quindi la distribuzione regolare identificata dalla funzione identicamente 1. Diremo, più semplicemente, che è la funzione 1:
$$\hat{\delta} = 1\,;$$
Quest'esempio mostra che la trasformata di una distribuzione che non è regolare può essere una distribuzione regolare.

5. Viceversa, sia $f(x) \equiv 1$. Dalla formula dell'antitrasformata si vede immediatamente che la sua trasformata è la distribuzione $2\pi\delta$.

   Indichiamo provvisoriamente con λ la distribuzione
$$\phi \to \langle\langle \lambda, \phi \rangle\rangle = -\phi'(0)$$
(vedremo più avanti un simboli migliore per identificare questa distribuzione). La sua trasformata di Fourier è
$$\phi \to \langle\langle \lambda, \hat{\phi} \rangle\rangle = -(\hat{\phi})'(\omega)_{|\omega=0} = \int_{-\infty}^{+\infty} ix\phi(x)\,\mathrm{d}x \ :$$
la sua trasformata è ix. Si noti:
$$\mathcal{F}(\lambda) = ix\mathcal{F}(\delta)\,. \ \blacksquare$$

Osservazione 401 Si è detto, nell'esempio 370 del paragrafo 4.1.3.1 che per indicare una distribuzione χ che agisce sulle funzioni ϕ della variabile x si indica anche la notazione $\chi(x)$, anche se ovviamente non ha alcun senso parlare del "valore in x" della distribuzione χ. Analogamente, per indicare $\mathcal{T}_h\chi$ è comune scrivere $\chi(x-h)$. Con queste notazioni si ha:

l	$\hat{l}$
1	$2\pi\delta(\omega)$
$e^{i\alpha x}$	$2\pi\delta(\omega-\alpha)$
δ	1
$\sin\alpha x$	$=-i\pi\{\delta(\omega-\alpha)-\delta(\omega+\alpha)\}$
$\cos\alpha x$	$\pi\{\delta(\omega-\alpha)+\delta(\omega+\alpha)\}$

4.3.4 Le operazioni sulle distribuzioni

Si sa già che $\mathcal{S}'$ è uno spazio lineare, ossia che in $\mathcal{S}'$ è definito il prodotto per scalari e la somma[10]; e in $\mathcal{S}'$ si è definita la trasformata di Fourier, con un metodo di trasposizione. Ancora con un metodo di trasposizione si definiscono altre operazioni, a partire dalle corrispondenti operazioni su $\mathcal{S}$.

1. Sia $h\in\mathbb{R}$ e $\mathcal{T}_h$ la traslazione in $\mathcal{S}$,

   $$(\mathcal{T}_h\phi)(x)=\phi(x-h)$$

   (se $h>0$ questa si interpreta come *traslazione verso destra*). Ovviamente, $\mathcal{T}_h$ è continua su $\mathcal{S}$ e quindi per ogni $l\in\mathcal{S}'$ si definisce

   $$\mathcal{T}_h'l\ :\quad \phi\to\langle\langle l,\mathcal{T}_h\phi\rangle\rangle\,.$$

   Se $h>0$ questa si interpreta come *traslazione (applicata ad l) verso sinistra.*

   Si esprima in modo esplicito l'effetto di $\mathcal{T}_h'$ sulle distribuzioni regolari e si giustifichi la notazione più comunemente usata $\mathcal{T}_{-h}$ invece di $\mathcal{T}_h'$.

[10] la continuità delle trasformazioni $\chi\to\alpha\chi$ e $\chi\to(\chi+\lambda)$ si dimostra facilmente.

Esempio 402 Sia δ la delta di Dirac. Si ha:

$$\langle\langle \mathcal{T}_h \delta, \phi \rangle\rangle = \langle\langle \delta, \mathcal{T}_{-h}\phi \rangle\rangle = \phi(h) = \langle\langle \delta(x-h), \phi \rangle\rangle \,. \quad \blacksquare$$

2. Sia ora $a \neq 0$ e sia

$$\mathcal{R}_a \;: \qquad (\mathcal{R}_a\phi)(x) = \phi(ax)\,.$$

Anche $\mathcal{R}_a$ è continua da $\mathcal{S}$ in sé e quindi si può definire $\mathcal{R}'_a$ ponendo

$$\langle\langle \mathcal{R}'_a l, \phi \rangle\rangle = \langle\langle l, \mathcal{R}_a\phi \rangle\rangle \,. \tag{4.26}$$

Si verifica facilmente che se D_f è la distribuzione regolare che corrisponde alla funzione integrabile f allora[11]

$$\mathcal{R}'_a(D_f) = D_{[\frac{1}{|a|}\mathcal{R}_{1/a} f]}$$

e quindi, per una generica $l \in \mathcal{S}'$, si definisce

$$\mathcal{R}_a l = \left[\frac{1}{|a|}\mathcal{R}_{1/a}\right]' l \quad \mathcal{R}'_a \text{ definita in (4.26}\,.$$

Ossia, $\mathcal{R}_a l$ è definita dall'uguaglianza

$$\langle\langle \mathcal{R}_a l, \phi \rangle\rangle = \langle\langle \left[\frac{1}{|a|}\mathcal{R}_{1/a}\right]' l, \phi \rangle\rangle = \langle\langle l, \frac{1}{|a|}\mathcal{R}_{1/a}\phi \rangle\rangle \,.$$

3. Si è notato che se $g \in C^\infty(\mathbb{R})$ è limitata con tutte le sue derivate, o anche se g è un polinomio, allora $\mathcal{M}_g$: $\mathcal{M}_g\phi = g\phi$ è continua da $\mathcal{S}$ in sé. Ciò permette di definire

$$\langle\langle (\mathcal{M}'_g l), \phi \rangle\rangle = \langle\langle l, \mathcal{M}_g\phi \rangle\rangle \,.$$

Si esamini l'azione di $\mathcal{M}'_g$ sulle distribuzioni regolari e si spieghi perchè si scrive

$$\mathcal{M}_g l$$

invece di $\mathcal{M}'_g l$.

Osservazione 403 La $\mathcal{M}_g$ è la "moltiplicazione" della *distribuzione* l per la *funzione* g. più comunemente, invece di scrivere $\mathcal{M}_g l$ si scrive gl. ∎

[11]si ricordi che in questa parte si lavora con funzioni e distribuzioni di una variabile. Si veda l'Esempio 362 per il caso generale.

Esempio 404 Sia $\psi(x) \in \mathcal{S}$. Si vuol calcolare $\mathcal{M}_\psi \delta_\alpha$. Si ha:

$$\langle\langle \mathcal{M}_\psi \delta_\alpha, \phi \rangle\rangle = \langle\langle \delta_\alpha, \psi\phi \rangle\rangle = \psi(\alpha)\phi(\alpha)\,.$$

Dunque, usando la notazione più comune $\psi(\omega)\delta(\omega-\alpha)$ invece di $\mathcal{M}_\psi \delta_\alpha$, si ha

$$\psi(\omega)\delta(\omega-\alpha) = \psi(\alpha)\delta(\omega-\alpha)\,. \tag{4.27}$$

Quindi, se $\psi(\alpha) = 0$ si ha $\psi(\omega)\delta(\omega-\alpha) = 0$. ∎

4. Introduciamo ora la DERIVATA delle distribuzioni.

   Essendo continua la trasformazione D

   $$\phi \to D\phi = \phi'$$

   da $\mathcal{S}$ in sé, definiremo

   $$\langle\langle D'l, \phi \rangle\rangle = \langle\langle l, D\phi \rangle\rangle\,.$$

   Si esamini l'effetto di D' sulle distribuzioni regolari e si chiarisca perché invece di D' si usa il simbolo $-D$. Con questa notazione Dl è definita da

   $$\langle\langle Dl, \phi \rangle\rangle = -\langle\langle l, D\phi \rangle\rangle\,.$$

   La derivata di distribuzioni si indica anche con l'apice:

   $$Dl = l'\,.$$

Esempio 405 Consideriamo la FUNZIONE DI HEAVISIDE

$$u(t) = \begin{cases} 0 & \text{se} \quad t < 0 \\ 1 & \text{se} \quad t > 0\,. \end{cases}$$

Questa funzione non è derivabile nel senso ordinario. E' però derivabile nel senso delle distribuzioni, ossia è derivabile la distribuzione regolare ad essa associata, e vale

$$\langle\langle DD_u, \phi \rangle\rangle = -\langle\langle u, \phi' \rangle\rangle = -\int_0^{+\infty} \phi'(s)\,\mathrm{d}s = \phi(0)\,.$$

Dunque la sua derivata è la delta di Dirac e scriveremo brevemente

$$Du = \delta\,.$$

Calcoliamo ora la derivata della delta di Dirac:

$$\langle\langle D\delta, \phi\rangle\rangle = -\langle\langle \delta, D\phi\rangle\rangle = -\phi'(0)\,.$$

Dunque, $D\delta$ è la trasformazione (che all'esempio 400 abbiamo chiamato $-\lambda$)

$$\phi \to -\phi'(0)\,. \blacksquare$$

Sia ora $f(x)$ una funzione di classe $C^1(\mathbb{R})$. Anche se $f(x)$ cresce al più come un polinomio, la sua derivata può non crescere polinomialmente. E quindi $D_{f'}$ può non identificare una distribuzione temperata. In tal caso la derivata di D_f, nel *senso delle distribuzioni temperate* non è $D_{f'}$, che non è una distribuzione temperata[12].

4.3.5 Operazioni e trasformata di Fourier

Studiamo ora le relazioni tra le operazioni introdotte in $\mathcal{S}'$ e la trasformata di Fourier.

1. Calcoliamo prima di tutto
$$\mathcal{F}(\mathcal{T}_h l)\,.$$
Per definizione,
$$\langle\langle \mathcal{F}(\mathcal{T}_h l), \phi\rangle\rangle = \langle\langle \mathcal{T}_h l, \hat{\phi}\rangle\rangle = \langle\langle l, \mathcal{T}_{-h}\hat{\phi}\rangle\rangle\,.$$
Ora,
$$\mathcal{T}_{-h}\hat{\phi} = \hat{\phi}(\omega+h) = \int_{-\infty}^{+\infty} e^{-i(\omega+h)x}\phi(x)\,\mathrm{d}x = \mathcal{F}(M_f\phi) \text{ ove } f(x) = e^{-ihx}\,.$$
Dunque,
$$\mathcal{F}(\mathcal{T}_h l) = M_f\hat{l} \ \text{ con } f(x) = e^{-ihx}\,.$$

2. Analogamente, la trasformata di $\mathcal{R}_a l$ si calcola da
$$\langle\langle \mathcal{F}(\mathcal{R}_a l), \phi\rangle\rangle = \langle\langle l, \frac{1}{|a|}\mathcal{R}_{1/a}\hat{\phi}\rangle\rangle\,.$$
Ora,
$$\left[\frac{1}{|a|}\mathcal{R}_{1/a}\hat{\phi}\right](\omega) = \frac{1}{|a|}\int_{-\infty}^{+\infty} e^{-i(\omega/a)x}\phi(x)\,\mathrm{d}x = \frac{1}{|a|}\int_{-\infty}^{+\infty} e^{-i\omega(x/a)}\phi(x)\,\mathrm{d}x =$$
$$\int_{-\infty}^{+\infty} e^{-i\omega t}\phi(at)\,\mathrm{d}t = \mathcal{F}(\mathcal{R}_a\phi)\,.$$

[12]ossia, la formula (4.4) vale in $\mathcal{S}'$ solo se anche f' definisce una distribuzione temperata.

Dunque,

$$\langle\langle \mathcal{F}(\mathcal{R}_a l), \phi\rangle\rangle = \langle\langle l, \mathcal{F}(\mathcal{R}_a \phi)\rangle\rangle = \langle\langle \mathcal{R}'_a \hat{l}, \phi\rangle\rangle = \langle\langle \frac{1}{|a|}\mathcal{R}_{1/a}\hat{l}, \phi\rangle\rangle \ :$$

$$\mathcal{F}(\mathcal{R}_a l) = \frac{1}{|a|}\mathcal{R}_{1/a}\hat{l}\,.$$

3. Veniamo infine alla trasformata della derivata di una distribuzione:

$$\langle\langle \mathcal{F}(Dl), \phi\rangle\rangle = \langle\langle Dl, \hat{\phi}\rangle\rangle = -\langle\langle l, D\hat{\phi}\rangle\rangle = -\langle\langle l, \mathcal{F}(\mathcal{M}_{-it}\phi)\rangle\rangle = -\langle\langle \mathcal{M}_{-it}\hat{l}, \phi\rangle\rangle$$

e quindi

$$\mathcal{F}(Dl) = M_f\hat{l} \text{ ove } f(\omega) = i\omega\,.$$

Si noti che in questo calcolo abbiamo usato la linearità di ambedue gli argomenti di $\langle\langle \cdot, \cdot\rangle\rangle$.

Osservazione 406 Si consideri la funzione di Heaviside $u(t)$. La sua derivata è δ e quindi

$$1 = \hat{\delta} = i\omega\hat{u}$$

da cui sembra di poter dedurre $\hat{u} = \frac{1}{i\omega}$. *Si noti che questa è una scrittura solamente formale,* a cui non abbiamo attribuito alcun significato, perché $1/(i\omega)$ non è integrabile.

Un calcolo diretto mostra

$$\begin{aligned}
&\langle\langle \hat{u}, \phi\rangle\rangle = \int_0^{+\infty} \hat{\phi}(\omega)\,\mathrm{d}\omega = \int_0^{+\infty}\left[\int_{-\infty}^{+\infty} e^{-i\omega x}\phi(x)\,\mathrm{d}x\right]\mathrm{d}\omega \\
&= \lim_{R\to+\infty}\int_0^R\left[\int_{-\infty}^{+\infty} e^{-i\omega x}\phi(x)\,\mathrm{d}x\right]\mathrm{d}\omega = \lim_{R\to+\infty}\int_{-\infty}^{+\infty}\left[\int_0^R e^{-i\omega x}\,\mathrm{d}\omega\right]\phi(x)\,\mathrm{d}x \\
&\lim_{R\to+\infty}\int_{-\infty}^{+\infty}\frac{1-e^{-iRx}}{ix}\phi(x)\,\mathrm{d}x\,.
\end{aligned}$$

Ciò mostra che l'azione di certe distribuzioni viene descritta mediante integrali dipendenti da parametri, e loro limiti. Noi non presentiamo questo aspetto del problema. Diciamo solamente che in questo modo si riesce a dare senso all'espressione $\hat{u} = 1/(i\omega)$. ∎

4.3.6 Convoluzione di distribuzioni

Ricordiamo che la convoluzione di due funzioni f e g è la funzione

$$(f * g)(t) = \int_{-\infty}^{+\infty} f(t-s)g(s)\,\mathrm{d}s$$

definita per quei valori di t per i quali l'integrale (nel senso di Lebesgue) converge.

Se f e g sono ambedue integrabili su $\mathbb{R}$ l'integrale converge q.o. e se f e g appartengono ad $\mathcal{S}$ l'integrale converge per ogni t.

Il problema di estendere il concetto di convoluzione al caso delle distribuzioni è piuttosto delicato, e ci limitiamo ad enunciare alcuni risultati. Consideriamo due funzioni integrabili f e g. La loro convoluzione $h = f * g$ identifica una distribuzione regolare D_h, la cui azione su $\phi \in \mathcal{S}$ è:

$$\int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} f(s-y)g(y)\,\mathrm{d}y\right] \phi(s)\,\mathrm{d}s = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} f(s-y)\phi(s)\,\mathrm{d}s\right] g(y)\,\mathrm{d}y$$
$$= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} f(x)\phi(x+y)\,\mathrm{d}x\right] g(y)\,\mathrm{d}y\,.$$

Per interpretare la formula precedente quando f e g sono sostituite da distribuzioni, è necessario considerare funzioni ϕ dipendenti da *due* variabili x ed y e, per ogni valore di una di esse, per esempio y, applicare una distribuzione l alla funzione $x \to \phi(x,y)$. Per intendere che l agisce sulla ϕ vista come funzione di x, scriveremo[13] l_x invece di l.

Siano ora l ed m due distribuzioni. Per definire il significato di $l * m$, CONVOLUZIONE di l e di m, dobbiamo spiegare come essa agisce su ciascuna funzione $\phi \in \mathcal{S}$. Per questo, scelta $\phi \in \mathcal{S}$, consideriamo la funzione $x \to \phi(x+y) = \mathcal{T}_{-y}\phi$, per ogni scelta di $y \in \mathbb{R}$, e consideriamo il numero[14]

$$\langle\!\langle m_x, \phi(x+y)\rangle\!\rangle = \langle\!\langle m_x, \mathcal{T}_{-y}\phi\rangle\!\rangle\,.$$

Si trova in questo modo una funzione

$$\psi(y) = \langle\!\langle m_x, \phi(x+y)\rangle\!\rangle\,.$$

Se accade che $\psi \in \mathcal{S}$, allora può definirsi

$$\langle\!\langle l, \psi\rangle\!\rangle = \langle\!\langle l_y, \psi\rangle\!\rangle$$

[13] abbiamo usato anche il simbolo $l(x)$. In questo paragrafo è probabilmente più chiaro usare il simbolo l_x che meno ricorda quello di una funzione della variabile x.

[14] ripetiamo che l'indice x di m_x indica che la funzione $\phi(x+y) = (\mathcal{T}_{-y}phi)(x)$ deve essere considerata una funzione di x per ogni valore del parametro y.

e per definizione porremo

$$\langle\langle l * m, \phi\rangle\rangle = \langle\langle (l * m)_x, \phi\rangle\rangle = \langle\langle l, \psi\rangle\rangle = \langle\langle l_y, \langle\langle m_x, \phi(x+y)\rangle\rangle\rangle\rangle$$
$$= \langle\langle l_y, \langle\langle m_x, \mathcal{T}_{-y}\phi\rangle\rangle\rangle\rangle \qquad (4.28)$$

dove $\mathcal{T}_{-y}\phi$ è la funzione $x \mapsto (\mathcal{T}_{-y}\phi)(x) = \phi(x+y)$.

Si noti che per poter utilizzare la definizione precedente abbiamo bisogno di più che non semplicemente $\psi \in \mathcal{S}$: abbiamo bisogno che la funzione

$$\phi(x) \to \psi(y) = \langle\langle m_x, \phi(x+y)\rangle\rangle$$

sia continua da $\mathcal{S}$ in sé, in modo da avere

$$\phi \to \langle\langle l * m, \phi\rangle\rangle$$

continua su $\mathcal{S}$.

Di conseguenza, la possibilità di definire la distribuzione $l*m$ dipende dalle proprietà delle distribuzioni con cui si lavora. I dettagli di quest'argomento sono piuttosto delicati e noi ci limitiamo ad enunciare i due risultati seguenti. Definiamo prima: una distribuzione l ha SUPPORTO in un insieme chiuso K quando $\langle\langle l, \phi\rangle\rangle = 0$ per ogni $\phi \in \mathcal{S}$ il cui supporto NON interseca K; ossia che è nulla in un aperto contenente K.

Vale:

Teorema 407 *La formula (4.28) definisce la convoluzione delle due distribuzioni temperate l ed m in uno dei due casi seguenti:*

- almeno una *delle due distribuzioni ha supporto compatto;*
- ambedue *le distribuzioni hanno supporto in* $[0, +\infty)$.

Inoltre:

Teorema 408 *Supponiamo che le distribuzioni che figurano nelle formule*

- *hanno tutte supporto compatto*

oppure

- *hanno tutte supporto in* $[0, +\infty)$.

In tal caso valgono le proprietà seguenti:

1. *distributività:*
$$l * (h + k) = l * h + l * k\,,$$
$$(l + m) * h = l * h + m * h\,;$$

2. *commutatività:*
$$h * k = k * h\,;$$

3. *associatività:*
$$l * (h * k) = (l * h) * k\,;$$

4. *regola di derivazione:*
$$D(l * k) = (Dl) * k = l * (Dk)\,;$$

5. *regola per la trasformata di Fourier:*
$$\mathcal{F}(l * h) = \hat{l}(\omega)\hat{h}(\omega)\,;$$

6. *esistenza dell'identità rispetto alla convoluzione:*
$$\delta * h = h\,.$$

La proprietà $\delta * h = h$ si interpreta dicendo che *la delta di Dirac è l'identità rispetto alla convoluzione.* L'Esempio 395 mostra che le identitá approssimate *approssimano la δ di Dirac,* ossia approssimano l'identità rispetto alla convoluzione. Si giustifica così il termine "identità approssimata".

Vediamo alcuni esempi:

Esempio 409 Per alleggerire la notazione, conviene usare f per indicare la distribuzione regolare D_f.

1. Sia $u(x)$ la funzione di Heaviside. La funzione è limitata e quindi identifica una distribuzione regolare con supporto $x \geq 0$. Dunque $u * u$ esiste nel senso delle distribuzioni e, come si calcola immediatamente anche nel senso integrale, perché
$$\int_{-\infty}^{+\infty} u(x - s)u(s)\,\mathrm{d}s = \int_0^x \mathrm{d}s = x\,.$$
Calcoliamo la convoluzione nel senso delle distribuzioni. Bisogna calcolare $\langle\langle u * u, \phi\rangle\rangle$ per ogni $\phi \in \mathcal{S}$:
$$\langle\langle u * u, \phi\rangle\rangle = \langle\langle (u * u)_x, \phi\rangle\rangle = \langle\langle u_y, \langle\langle u_x, \mathcal{T}_{-y}\phi\rangle\rangle\rangle\rangle$$

E':

$$\langle\langle u_x, \mathcal{T}_{-y}\phi\rangle\rangle = \int_0^{+\infty} \phi(x+y)\,\mathrm{d}x$$

e quindi

$$\begin{aligned}\langle\langle u_y, \langle\langle u_x, \mathcal{T}_{-y}\phi\rangle\rangle\rangle\rangle &= \int_0^{+\infty}\int_0^{+\infty} \phi(x+y)\,\mathrm{d}x\;\mathrm{d}y \\ = \int_0^{+\infty}\int_y^{+\infty}\phi(s)\,\mathrm{d}s\;\mathrm{d}y &= \int_0^{+\infty}\int_0^{s}\phi(s)\,\mathrm{d}y\,\mathrm{d}s = \int_0^{+\infty} s\phi(s)\,\mathrm{d}s\,.\end{aligned}$$

Dunque la distribuzione $u * u$ è la distribuzione regolare $f(x) = x$, in accordo col risultato trovato calcolando la convoluzione delle funzioni.

2. Sia ancora u la distribuzione regolare identificata dalla funzione di Heaviside e calcoliamo $(D^k\delta) * u$ per ogni k intero non negativo (se $k = 0$ intendiamo $D^0\delta = \delta$). Dunque calcoliamo

$$\langle\langle (D^k\delta * u)_x, \phi\rangle\rangle = \langle\langle (D^k\delta)_y, \langle\langle u_x, \mathcal{T}_{-y}\phi\rangle\rangle\rangle\rangle\,.$$

E'

$$\langle\langle u_x, \mathcal{T}_{-y}\phi\rangle\rangle = \int_0^{+\infty} \phi(x+y)\,\mathrm{d}y$$

e quindi

$$\begin{cases} \text{se } k>0 \text{ si ha} \\ \langle\langle (D^k\delta)_y, \langle\langle u_x, \mathcal{T}_{-y}\phi\rangle\rangle\rangle\rangle = (-1)^k \displaystyle\int_0^{+\infty} \phi^{(k)}(x)\,\mathrm{d}x = (-1)^{k-1}\phi^{(k-1)}(0) \end{cases}$$
$$\begin{cases} \text{se } k=0 \text{ si ha} \\ \langle\langle (D^k\delta)_y, \langle\langle u_x, \mathcal{T}_{-y}\phi\rangle\rangle\rangle\rangle = \displaystyle\int_0^{+\infty} \phi(x)\,\mathrm{d}x\,. \end{cases}$$

Dunque,

$$(D^k\delta)*u = \begin{cases} (-1)^{k-1}D^{k-1}\delta & \text{se} \quad k>0 \text{ (se } k-1=0 \text{ si intende } D^0\delta=\delta) \\ u & \text{se} \quad k=0 \text{ (ove } u \text{ è la funzione di Heaviside)}\,. \end{cases}$$

3. Se $k = 0$ poniamo $D^0 l = l$ e in particolare $D^0\delta = \delta$. Sia k intero non negativo e calcoliamo $l * D^k\delta$:

$$\langle\!\langle (l * D^k\delta)_x, \phi \rangle\!\rangle = \langle\!\langle l_y, \langle\!\langle (D^k\delta)_x, \mathcal{T}_{-y}\phi \rangle\!\rangle \rangle\!\rangle$$
$$= \langle\!\langle l_y, \langle\!\langle \delta_x, (-1)^k (D^k \mathcal{T}_{-y}\phi) \rangle\!\rangle \rangle\!\rangle = \langle\!\langle l_y, (-1)^k D^k\phi \rangle\!\rangle$$

e quindi $l * D^k\delta = D^k l$.

In particolare, $l * \delta = l$.

Se nel calcolo precedente l è una distribuzione regolare f, rappresentata da una funzione k volte derivabile allora

$$f * D^k\delta = f^{(k)} .$$

Se in particolare $f(x) \equiv 1$ allora $f * D^k\delta = 0$ per ogni $k > 0$.

4. Le condizioni del Teorema 407 sono solo *sufficienti* per avere che $y \mapsto \langle\!\langle m_x, \mathcal{T}_{-y}\phi \rangle\!\rangle$ sia un elemento di $\mathcal{S}$ e quindi per poter definire $l * m$. Se la convoluzione si può definire senza che valgano le condizioni del Teorema 407 allora possono verificarsi fatti che rendono la convoluzione di scarsa utilità.

   Consideriamo gli esempi seguenti:

   (a) Torniamo a considerare l'esempio 3. La distribuzione l di quest'esempio è un qualsiasi elemento di $\mathcal{S}$, e non deve soddisfare le condizioni del Teorema 407. Se tali condizioni valgono allora

   $$y \mapsto \langle\!\langle l_x, \phi(x+y) \rangle\!\rangle \in \mathcal{S} \text{ e si può calcolare } \langle\!\langle m_y, \langle\!\langle l_x, \phi(x+y) \rangle\!\rangle \rangle\!\rangle$$

   per ogni distribuzione temperata m. Consideriamo però la distribuzione temperata $l = \mathbb{1}$ che è la distribuzione regolare identificata dalla funzione identicamente uguale ad 1, $\mathbb{1}(x) = 1$ per ogni $x \in \mathbb{R}$. In questo caso,

   $$\langle\!\langle \mathbb{1}_x, \mathcal{T}_{-y}\phi \rangle\!\rangle = \int_{-\infty}^{+\infty} \phi(x+y)\,\mathrm{d}x = \int_{-\infty}^{+\infty} \phi(x)\,\mathrm{d}x$$

   è costante e quindi non appartiene ad $\mathcal{S}$. In particolare per esempio non esiste $\langle\!\langle u_y, \langle\!\langle l_x, \phi(x+y) \rangle\!\rangle \rangle\!\rangle$ quando u è la funzione di Heaviside; ossia, $u * \mathbb{1}$ non esiste. Usando il fatto che le funzioni costanti sono in particolare distribuzioni temperate regolari e il fatto che il supporto di $D^k\delta$ è $\{0\}$ è invece possibile dare un senso a $D^k\delta * l$ e inoltre, si

può provare che $D^k\delta * l = D^k l = l * D^k\delta$. In particolare, se $k = 0$ si tottiene $\delta * l = l * \delta = l$.

Quindi, si può attribuire un significato al simbolo $m * \mathbb{1}$ per certe distribuzioni temperate m e non per altre.

(b) Mostriamo ora che se non valgono le condizioni del Teorema 407 generalmente non vale nemmeno la proprietà associativa. Consideriamo le tre distribuzioni seguenti: $\mathbb{1}$ è ancora la distribuzione regolare identificata dalla funzione identicamente uguale ad 1, $\mathbb{1}(x) = 1$ per ogni $x \in \mathbb{R}$; $u(x)$ è la distribuzione regolare identificata dalla funzione di Heaviside e la terza distribuzione è δ'.

Calcoliamo la distribuzione $(\mathbb{1} * \delta') * u$ e la distribuzione $\mathbb{1} * (\delta' * u)$. La prima è nulla,

$$(\mathbb{1} * \delta') * u = 0$$

perché, come si è calcolato nell'esempio 3, $\mathbb{1} * \delta' = 0$.

Usando nuovamente l'esempio 3, calcoliamo $\mathbb{1} * (\delta' * u)$ e notiamo che non è nulla. Infatti, $\delta' * u = \delta$ e quindi

$$\mathbb{1} * (\delta' * u) = \mathbb{1} * \delta = \mathbb{1}\,.$$

Dunque,

$$\mathbb{1} * (\delta' * u) = \mathbb{1} * \delta = \mathbb{1} \neq 0 = (\mathbb{1} * \delta') * u)\,.$$

4.3.7 Treno d'impulsi e formula di Poisson

Indichiamo, per ora in modo formale, col simbolo L la formula

$$L = \sum_{n=-\infty}^{+\infty} \delta(x - n)\,.$$

Intendiamo con questo la distribuzione che a ϕ associa

$$\langle\langle L, \phi \rangle\rangle = \sum_{n=-\infty}^{+\infty} \phi(n)\,. \tag{4.29}$$

Ovviamente dobbiamo ancora provare che L effettivamente definisce un elemento di $\mathcal{S}'$. Lo proveremo in due modi diversi, e la distribuzione temperata L si chiama TRENO D'IMPULSI. In modo più immaginoso, talvolta viene anche chiamata "pettine di Dirac".

Mostriamo ora che effettivamente $L \in \mathcal{S}'$ basandoci solamente sulla definizione di $\mathcal{S}'$. Essendo ovviamente L lineare, basta provarne la continuità; ossia va provato che se (ϕ_k) converge a zero in $\mathcal{S}$ allora la successione $(L\phi_k)$ converge a zero in $\mathbb{R}$.

Ricordiamo che se (ϕ_k) converge a zero in $\mathcal{S}$ allora in particolare le successioni di funzioni $(\phi_k(x))$ e $(x^2\phi_k(x))$ convergono a zero uniformemente su $\mathbb{R}$. Dunque, per ogni $\epsilon > 0$ esiste K_ϵ tale che per ogni $k > K_\epsilon$ ed ogni $x \in \mathbb{R}$ si ha

$$|\phi_k(x)| < \epsilon\,, \qquad |x^2\phi_k(x)| < \epsilon\,.\ \text{In particolare } |\phi_k(n)| \le \frac{\epsilon}{n^2} \text{ per ogni } n\,.$$

Usando queste disuguaglianze si vede che per $k > K_\epsilon$ si ha

$$|\langle\langle L, \phi_k\rangle\rangle| = \left|\sum_{n=-\infty}^{+\infty} \phi_k(n)\right| \le |\phi_k(0)| + \sum_{n\neq 0} \frac{\epsilon}{n^2} \le \epsilon\left\{1 + \sum_{n\neq 0}\frac{1}{n^2}\right\} < (\text{const})\cdot\epsilon\,.$$

Quindi, il treno di impulsi è una distribuzione temperata, come volevamo provare.

Una via un po' più involuta per arrivare allo stesso risultato è la seguente. Conviene vederla perché condurrà ad una uguaglianza interessante.

Consideriamo l'estensione periodica, di periodo 1, della funzione

(se vogliamo una funzione definita su $\mathbb{R}$ si può anche imporre $f(0) = 0$). Usiamo ancora la notazione $f(x)$ per indicare l'estensione. Dunque:

$$\begin{aligned} &\text{se} \quad n < x < n+1 \quad \text{allora} \\ &f(x) = f(x-(n+1)) = x-(n+1)+\frac{1}{2} = x-n-\frac{1}{2}\,. \end{aligned} \tag{4.30}$$

Il grafico della funzione è nella figura 4.2. Essendo $f(x)$ limitata su $\mathbb{R}$, essa definisce una distribuzione temperata, che ancora denotiamo f, la cui derivata si calcola immediatamente: è

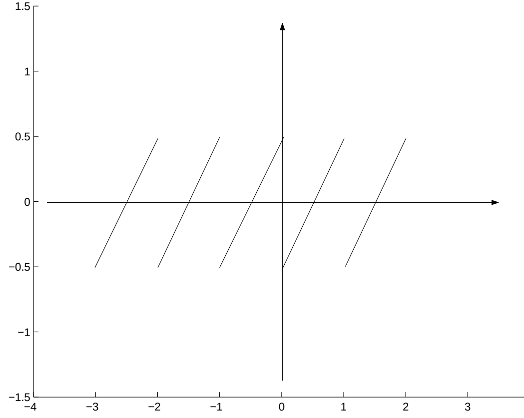
$$Df = 1 - \sum_{n=-\infty}^{+\infty}\delta(x-n) \text{ e quindi } \sum_{n=-\infty}^{+\infty}\delta(x-n) = 1 - Df$$

e anche per questa via si vede che il treno d'impulsi è una distribuzione temperata.

Essendo L una distribuzione temperata, ha una trasformata di Fourier. La continuità della trasformata di Fourier mostra che

$$\mathcal{F}\left(\sum_{n=-\infty}^{+\infty}\delta(x-n)\right) = \sum_{n=-\infty}^{+\infty} e^{-i\omega n} = \sum_{n=-\infty}^{+\infty} e^{i\omega n} = 1+2\left(\sum_{n=1}^{+\infty}\cos\omega n\right)\,. \tag{4.31}$$

Figura 4.2:



In particolare ciò mostra che le serie di funzioni (ossia di distribuzioni regolari!) a secondo membro convergono in $\mathcal{S}'$ nonostante che le serie non convergano puntualmente.

Vogliamo ora trovare un'espressione diversa per la trasformata.

Sviluppiamo la funzione $f(x)$ in (4.30) in serie di Fourier. Si ottiene

$$f(x) = -\frac{1}{\pi}\sum_{n=1}^{+\infty}\frac{1}{n}\sin(2n\pi x)\,.$$

Questa serie converge in $L^2(-n,n)$ per ogni n. Mostriamo che essa converge anche nel senso di $\mathcal{S}'$. Ignorando il fattore $-1/\pi$, consideriamo, per ogni $\phi\in\mathcal{S}$,

$$\begin{aligned}
&\lim_N \langle\langle \sum_{n=1}^{N}\frac{1}{n}\sin(2n\pi x),\phi\rangle\rangle = \lim_N\int_{-\infty}^{+\infty}\left[\sum_{n=1}^{N}\frac{1}{n}\sin(2n\pi x)\right]\phi(x)\,\mathrm{d}x\\
&= -\lim_N\int_{-\infty}^{+\infty}\frac{\mathrm{d}}{\mathrm{d}x}\left[\sum_{n=1}^{N}\frac{1}{2n^2\pi}\cos(2n\pi x)\right]\phi(x)\,\mathrm{d}x\\
&= \lim_N\int_{-\infty}^{+\infty}\left[\sum_{n=1}^{N}\frac{1}{2n^2\pi}\cos(2n\pi x)\right]\phi'(x)\,\mathrm{d}x\\
&= \int_{-\infty}^{+\infty}\left[\sum_{n=1}^{+\infty}\frac{1}{2n^2\pi}\cos(2n\pi x)\right]\phi'(x)\,\mathrm{d}x
\end{aligned}$$

grazie al teorema della convergenza dominata. Dunque, la serie converge in $\mathcal{S}'$. La continuità della trasformata di Fourier mostra che la trasformata di $f(x)$ può calcolarsi scambiando la trasformata col segno di serie:

$$\hat{f}(\omega) = i\sum_{n=1}^{+\infty} \frac{1}{n}\left[\delta(\omega - 2n\pi) - \delta(\omega + 2n\pi)\right] = i\sum_{n\neq 0} \frac{1}{n}\delta(\omega - 2n\pi)$$
$$= -i\sum_{n\neq 0} \frac{1}{n}\delta(\omega + 2n\pi)\,.$$

Dunque, la trasformata del treno di impulsi è anche[15]

$$\mathcal{F}\left(\sum_{n=-\infty}^{+\infty} \delta(x-n)\right) = \mathcal{F}\left(1 - Df\right) = 2\pi\delta(\omega) - i\omega\hat{f}(\omega)$$
$$= 2\pi\delta(\omega) + \sum_{n\neq 0} \frac{\omega}{n}\delta(\omega + 2n\pi) = 2\pi\delta(\omega) + \sum_{n\neq 0} \frac{2n\pi}{n}\delta(\omega + 2n\pi)$$
$$= 2\pi \sum_{n=-\infty}^{+\infty} \delta(\omega - 2n\pi)\,. \tag{4.32}$$

Questa è la forma più comunemente usata per la trasformata del treno d'impulsi. Combinandola con la rappresentazione (4.31) si trova

$$2\pi \sum_{n=-\infty}^{+\infty} \delta(\omega - 2n\pi) = \sum_{n=-\infty}^{+\infty} e^{-i\omega n} = \sum_{n=-\infty}^{+\infty} e^{i\omega n} = 1 + 2\left(\sum_{n=1}^{+\infty} \cos \omega n\right)\,.$$

L'uguaglianza va intesa come uguaglianza in $\mathcal{S}'$; ossia, applicando i due membri ad una $\phi \in \mathcal{S}$ si trova lo stesso valore. Si ha quindi

$$2\pi\langle\!\langle \sum_{n=-\infty}^{+\infty} \delta(\omega - 2n\pi), \phi\rangle\!\rangle = 2\pi \sum_{n=-\infty}^{+\infty} \langle\!\langle\delta(\omega - 2n\pi), \phi\rangle\!\rangle$$
$$= 2\pi \sum_{n=-\infty}^{+\infty} \phi(2n\pi) = \langle\!\langle \sum_{n=-\infty}^{+\infty} e^{i\omega n}, \phi\rangle\!\rangle = \sum_{n=-\infty}^{+\infty} \langle\!\langle e^{i\omega n}, \phi\rangle\!\rangle$$
$$= \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{i\omega n}\phi(\omega)\,\mathrm{d}\omega = \sum_{n=-\infty}^{+\infty} \hat{\phi}(-n) = \sum_{n=-\infty}^{+\infty} \hat{\phi}(n)\,.$$

[15] si ricordi che:1) la notazione $\delta(x)$ si usa per indicare che la distribuzione agisce su funzioni della variabile x; 2) la notazione $\delta(x-a)$ si usa per indicare la distribuzione δ_a; 3) per ogni polinomio vale $p(x)\delta(x-a) = p(a)\delta(x-a)$, si veda la (4.27).

Si è così trovata l'uguaglianza seguente, nota come FORMULA DI POISSON:

$$2\pi \sum_{n=-\infty}^{+\infty} \phi(2\pi n) = \sum_{n=-\infty}^{+\infty} \hat{\phi}(n) \qquad \forall \phi \in \mathcal{S}\,.$$

4.3.8 Avvertenza finale e ricapitolazione

Abbiamo trattato la trasformata di Fourier per il caso delle funzioni che dipendono da *una sola* variabile. Comesi è notato, in pratica è necessario lavorare anche con la trasformata di Fourier di funzioni che dipendono da n variabili $x_1, \dots, x_n$, $n > 1$. Si è già detto che in questo caso la trasformata di Fourier è la funzione *ancora di n variabili* $\xi_1, \dots, \xi_n$. La trasformata di Fourier di $f(x_1, \dots, x_n)$ è

$$\begin{aligned}\hat{f}(\xi_1, \dots, \xi_n) &= \int_{\mathbb{R}^n} e^{-i\xi\cdot x} f(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^n} e^{-i[x_1\xi_1+\cdots+x_n\xi_n]} f(x_1, \dots, x_n)\,\mathrm{d}x_1 \dots \mathrm{d}x_n\,.\end{aligned}$$

Praticamente nessun cambiamento va fatto a ciò che abbiamo visto per le funzioni di una variabile, salvo che la formula per l'antitrasformata è ora

$$f(x_1, \dots, x_n) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i[x_1\xi_1+\cdots+x_n\xi_n]} \hat{f}(\xi_1, \dots, \xi_n)\,\mathrm{d}\xi_1 \dots \mathrm{d}\xi_n\,.$$

Naturalmente questa formula vale sotto ipotesi alquanto restrittive ma, esattamente come nel caso delle funzioni di una sola variabile, essa si estende al caso delle distribuzioni.

E' importante notare che testi diversi usano definizioni lievemente diverse della trasformata di Fourier. In particolare, nel caso di funzioni di una variabile,

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} e^{-2\pi i\omega t} f(t)\,\mathrm{d}t \qquad \text{oppure} \qquad \hat{f}(i\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-i\omega t} f(t)\,\mathrm{d}t\,.$$

Ovviamente questo non altera la teoria che abbiamo presentato, mentre le formule vengono lievemente diverse: differiscono per un fattore a causa del fatto che le formule di inversione sono, rispettivamente,

$$f(t) = \int_{-\infty}^{+\infty} e^{+2\pi i\omega t} \hat{f}(\omega)\,\mathrm{d}\omega \qquad \text{oppure} \qquad f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{+i\omega t} f(t)\,\mathrm{d}\,.$$

Più raramente si trova la definizione

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} e^{+i\omega t} f(t)\,\mathrm{d}t\,.$$

Infine raccogliamo gli esempi di trasformate di Fourier che abbiamo visto:

$f(t)$	$\hat{f}(\omega)$
$e^{-\|t\|}$	$\frac{2}{1+\omega^2}$
$\chi_{[-T,T]}(t)$	$\frac{\sin \omega T}{\omega}$
$e^{-x^2/2}$	$\sqrt{2\pi} e^{-\omega^2/2}$
$e^{i\alpha t}$	$2\pi\delta_\alpha$
$\sin \alpha x$	$-i\pi\{\delta_\alpha - \delta_{-\alpha}\}$
$\cos \alpha x$	$\pi\{\delta_\alpha + \delta_{-\alpha}\}$
δ	1
1	$2\pi\delta$
Heaviside	$\phi \to \int_0^{+\infty} \hat{\phi}(\omega)\,\mathrm{d}\omega$

Indice analitico